https://doi.org/10.24191/jeesr.v21i1.014

MSVM Modelling on Agarwood Oil Various Qualities Classification

Aqib Fawwaz Mohd Amidon, Noratikah Zawani Mahabob, Siti Mariatul Hazwa Mohd Huzir, Zakiah Mohd Yusoff*, Nurlaila Ismail, and Mohd Nasir Taib

***Abstract*—The aroma of agarwood oil is frequently connected with wealth and prestige. Because of its aroma, the oil is in great demand all over the world. Unfortunately, no standard method for classifying the quality of agarwood oil has been proposed to the world. This is essential to preserving the quality of agarwood oil and preventing fraudulent quality in the market. With that, for agarwood oil grading, a quality classification model must be developed. As part of the ongoing study for this standard's model development, an intelligent algorithm function has been implemented to ensure the model's capability is completely unquestionable. The intelligent algorithm used was Support Vector Machine (SVM) as main structure model. Then, the model has been added with Multiclass Classifier function by using One Verses One (OVO) strategy. The previous researcher's data was used in the analysis process, which included four classes of agarwood oil quality samples: low, medium low, medium high, and high quality. The result was a quality categorization of low, medium low, medium high or high quality, whereas the input was chemical abundances (percentages). The combination of both functions produced extremely good results for agarwood oil quality classification. The simulation platform was MATLAB software version r2020a, which was used for the desk study. The findings of this study will undoubtedly be useful in future agarwood oil research, particularly in the grade categorization section.**

***Index Terms*— Agarwood oil, Classification, Support Vector Machine, Multiclass Classifier, One Versus One Strategy.**

I. INTRODUCTION

AGARWOOD, also known as "Black Gold," is a highly valuable aromatic dark resinous Aquilaria species heartwood [1], [2]. It is also known as agar in India, jinkoh in Japan, chenxiang in China, oud in the Middle East, and gaharu in Southeast Asia [2]-[7]. The wounding and fungal infection of Aquilaria trees is the way to produce agarwood. As a defensive mechanism, trees exude resin, which is then deposited around wounds over time, eventually becoming agarwood [8]. For generations, agarwood has been used for medicinal scents, traditional medicine, religious ceremonies, and as a fragrant culinary component [1]-[7]. Agarwood keeps becoming one of the most in-demand products to be implemented as a raw material for incense, perfume, and medicine all over the world, with the Middle East and East Asia being the two most active markets with high demand [2], [4], [6], [9].

In recent decades, the demand for supplying agarwood oil has been increasing in the consuming countries. According to their quality, agarwood values range from 20usd to 6,000usd per kilograms for the wood itself [10]. Also, the price of agarwood oil per kilogram can achieve until 30,000 usd. The global market is projected to be worth between 6usd to 8usd billion per year [11].

Agarwood oil is usually distributed in varying quality. There is no standard classification model of agarwood oil introduced until now. All the producer countries used their own method to classify the quality of agarwood oil [12],[13]. Agarwood oil is divided into some grades in Malaysia and labelled as A, B, C, and D [14],[15]. In India, agarwood oil is classified into four classes using numbering which are 1,2,3, and 4 based on its color appearance [16]. While in Japan, agarwood oil is categorized as either high or low quality, with Kanankoh being the highest and Jinkoh being the lowest grade. But in Malaysians they classified Kalambak as high grade and gaharu as low grade [17]. This method applied in certain area of Malaysia. This proves that agarwood oil categorization varies according to countries and the grading method is not standardized.

According to a review of previous research studies [18]–[20], only human sensory panels have been used to grade agarwood oil. Producer countries hire experts to differentiate between the high and low quality of agarwood oil. They use their eyes to see the concentration of color in each of the samples. Then, they use their nose to evaluate the level of odor produced by each of the samples. This practice, however, may create concerns since it is ineffectual. In order to classify the huge number of agarwood oil samples using the current method, it certainly has an impact on those involved in classifying the quality of the agarwood oil [18]. Moreover, the classification process requires a long period of time because of the limitation of the ability of the human nose, which is not able to smell odors in large amounts over a long period of time. In addition, the cost of financing the classification process is also quite high. This is because it is necessary to consider the payment to the expert,

This manuscript is submitted on 07th February 2022 and accepted on 28 July 2022.

Aqib Fawwaz Mohd Amidon, Noratikah Zawawi Mahabob and Nurlaila Ismail are with the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor whereas Siti Mariatul Hazwa Mohd Huzir and Zakiah Mohd Yusoff are with campus of Pasir Gudang, Johor Branch. Mohd Nasir Taib is from MITRANS, Universiti Teknologi MARA, 40450 Shah Alam, Selangor

*Corresponding author
Email address: zakiah9018@uitm.edu.my

the financing of the training period to train new experts, production process and so on [19],[20].

In the development of studies on agarwood today, there is platform where agarwood oil quality classification can be done purely based on their chemical compounds, allowing essential oils to be classed and measured according to their different classes such as high, medium high, medium low or low quality, because of data analysis technologies todays. Intelligent technique designed for classification data such as Artificial Neural Network (ANN). K-Nearest Neighbors (KNN), Self-Organizing Map (SOM), Multilayer Perceptron (MLP), Support Vector Machine (SVM) and Radial Basis Function (RBF) are the examples that have been used to classify agarwood oil quality [21]-[23]. All the intelligent techniques mentioned were developed to classify the data. Some of the techniques were only able to classify linear data, and some of them were able to classify both linear and non-linear data [21],[23]. However, in this study, the Support Vector Machine algorithm has been selected to be used because SVM works relatively excellent when there is a clear margin of separation between two different classes and also the algorithm is relatively memory efficient compared to others.

This study employs the Support Vector Machine as a major model structure. Eventhough, the SVM technique has a limitation. This because, the technique was developed to classify two classes, not more than that as shown in Fig. 1, but it can be improved with supported by Multiclass Classifier as an algorithm function. There are three techniques or known as strategy to deploy these Multiclass Classifier machine learning algorithms. The first strategy is known as One Versus One (OVO) strategy. The second alternative is One Versus All (OVA) strategy. Error Correcting Output Codes (ECOC) is the third approach [24]-[26]. With that, Multiclass Classifier have been implemented by using One Versus One (OVO) strategy This is because the OVO strategy divides the dataset into one dataset for each class versus every other class, as opposed to the OVA approach, which separates it into one binary dataset for each class, as shown in Fig. 2. Plus, ECOC approach also not really recommended because it builds several subordinate classifiers on its own, each one responsible for removing some doubt about the proper input class. Then, using a voting system, determine which class is accurate. Unlike the OVO approach, it classifies each class versus class, class by class.

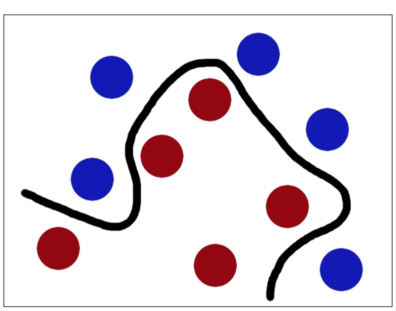

Fig. 1. The concept of Support Vector Machine technique classification for binary class of nonlinear data.

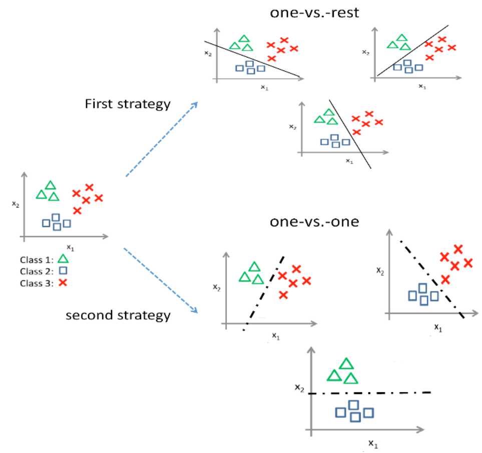

Fig. 2. The comparison of OVO and OVA approach in classifying multiclass of dataset.

II. METHODOLOGY

Throughout the research, four different agarwood oil grade samples totaling 660 were employed. As indicated in Table 1, these data samples have been sorted into 330 samples of good quality data, 30 samples of medium high quality data, 90 samples of medium low quality, and 210 samples of low quality data. All the data samples have eleven selected significant compounds as recommended by [1],[3].

TABLE I
DATA SAMPLES BASED ON THE GRADES

Grades	Quantity	Percent(%)
High	330	50.00
MediumHigh	30	4.55
MediumLow	90	13.64
Low	210	31.81
Total	660	100.00

A. Experimental Set – Up

Fig. 3 explain the details about the Multiclass Support Vector Machine development. First, All of agarwood data have gone through the data pre- processing stage, which includes normalizing data, randomizing data arrangement, and lastly separating training and testing data using an 80:20 percent ratio as proposed by previous studies [1], as shown in Table 2. Next, in the simulation, the Holdout partition was employed to isolate data, as proposed [3].

Following that, Support Vector Machine (SVM) model was developed by modifying a conventional template design for nonlinear data. By having the capability to analyze nonlinear data, SVM have been employed as a model, classification as the way of analysis of model, and 'gaussian' as the kernel function parameter to analyze nonlinear data. Next, support vector values were needed in order to make the classification process successful. Support vectors are data points that are closer to the hyperplane and have an impact on the hyperplane's location and orientation. Maximization of the classifier's margin can be done by using these support vectors. The location of the hyperplane will vary if the support vectors are deleted.Because the SVM technique cannot be used to create four different grading classifications simultaneously at one time, the multiclass classification approach based on binary classifiers was then utilized to ensure that the classification model was successful. The One versus One (OVO) strategy was chosen as the multiclass approach of machine learning in the multiclass classification. Then, each of quality have been classified as Low vs High, Low vs Medium High, Low vs Medium Low, Medium Low vs High, Medium Low vs Medium High and Medium High vs High.

Then, all 528 of the training data were employed in the development of the SVM model. Initially, 132 samples of testing data were maintained in order to test and evaluate the model using the performance criteria as a standard pass assessment for the SVM model. The testing findings have been presented in the results and discussion section.

To categorize four different qualities of agarwood oil, the simulation of an intelligent model was carried out using Matlab software version r2020a.

TABLE II
RATIO BETWEEN TRAINING DATA AND TESTING DATA

	No.ofSamples	Ratio	Percent(%)
TrainingData	528	4:1	80
TestingData	132		20

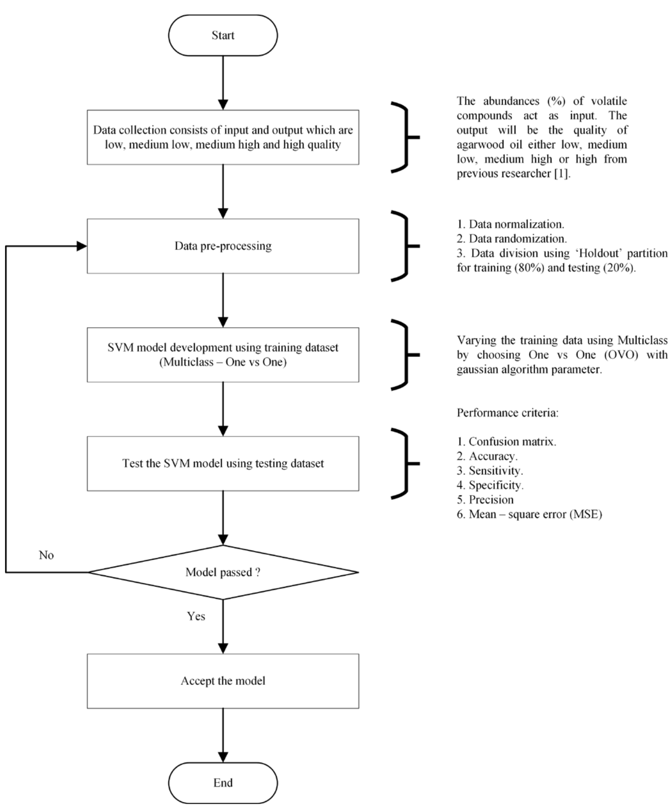

Fig. 3. Flowchart of detail experiment of Multiclass Support Vector Machine Based on One Versus One Strategies.

III. RESULTS AND DISCUSSION

The findings were separated into two sections; (i) Generate support vectors on classification process and (ii) Performance criteria of MSVM One Versus One Strategies evaluation.

A. Generate Support Vectors on Classification Process

The execution of the MSVM One Versus One (OVO) strategy resulted in varying positions of support vectors, as shown in Fig. 4 until Fig. 9, depending on the type of qualities being compared. Next, the function of training data set is to complete the SVM model in order to yield support vectors. The four different classes of agarwood oil qualities' support vector have been separated by hyperplane during one versus one classification process.

Overall, the results obtained show that the theory of the OVO strategy is correct. All Figures 4 to 9 show that this OVO strategy classifies classes vs classes, one by one. This makes it easier to analyse each classification process. Fig. 4 shows the good classification difference between low quality agarwood oil and high quality agarwood oil. This is evidenced when all the resulting support vector values are from low quality and high quality group data only after the SVM technique is implemented. Certainly, it is easy for the MSVM classification model to make a classification because the abundance of chemical compounds in high quality group is much higher and much better than the low qualitygroup.

Similarly, as seen in Figures 5 to 9, which refers to the outcome of support vectors value position of low vs medium high quality, low vs medium low quality, medium low vs high quality, medium low vs medium high quality. All of the support vector value positions were located in the comparing group area as mentioned before. Figures 6, 7, and 9 are also seen to produce support vectors only from the quality groups being compared. There are a few differences in Figures 5 and 8. In Fig. 5, the comparison conducted between the low quality group and the medium high group, but there are also support vector values from the resulting high quality group. Similarly, in Fig. 8, the comparison made is between medium low and medium high quality, but there is also a support vector value from the high quality group. This indicates that there are several datasets from the high quality group that seem to have almost the same abundance of significant chemical compounds as the medium high quality group.

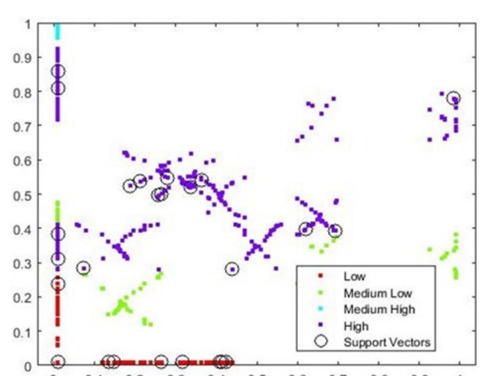

Fig. 4. Support vectors position for Low vs High quality.

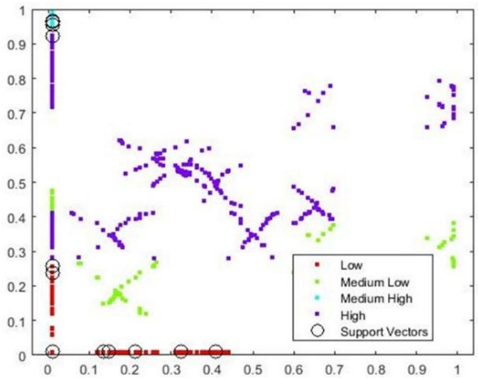

Fig. 5. Support vectors position for Low vs Medium High quality.

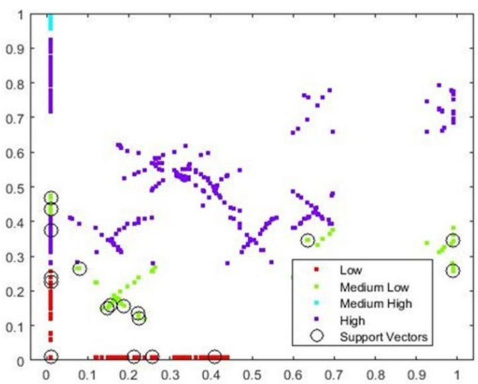

Fig. 6. Support vectors position for Low vs Medium Low quality.

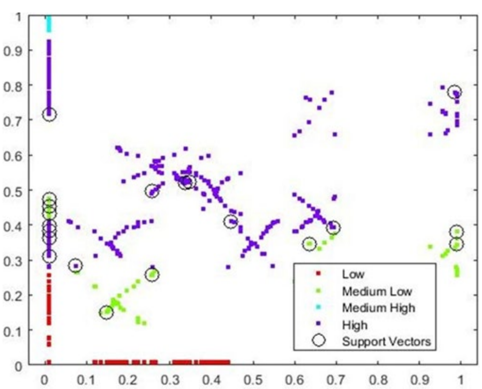

Fig. 7. Support vectors for Medium Low vs High quality.

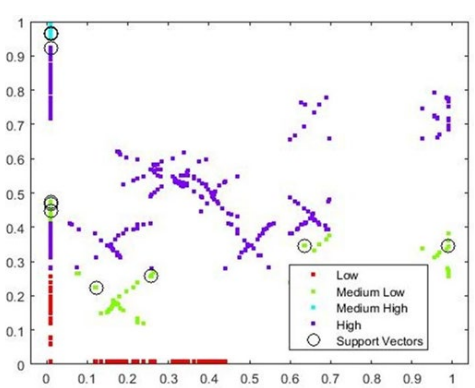

Fig. 8. Support vectors position for Medium Low vs Medium High quality.

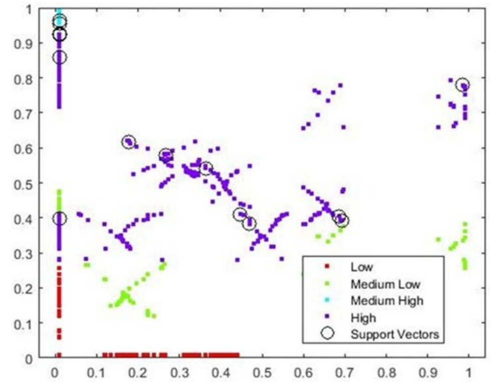

Fig. 9. Support vectors position for Medium High vs High quality.

Fig. 10 depicts a summary of the support vectors' values based on the OVO grade of quality. The data that is closest to the hyperplane is the support vectors. Low vs High has larger supportvector values than the others comparing. Thissignifies that Low vs High maximized the distance between the support vectors and the hyperplane in order to have higher support vectors during classification process occur.

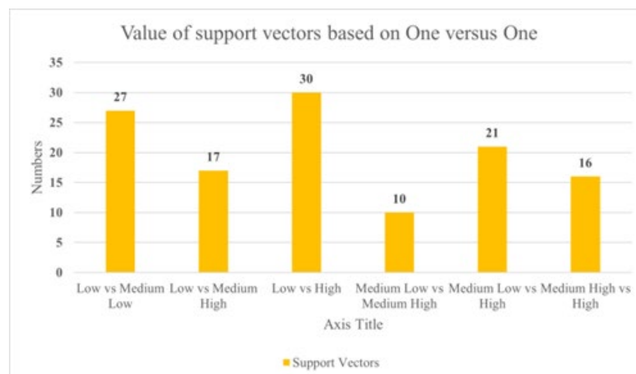

Fig. 10. Summarization of support vectors quality based on One Versus One Strategy.

B. *Performance criteria of MSVM One Versus One strategy evaluation*

Fig. 11 demonstrates that there is no mismatch between actual and predicted quality categorization. As a result, the error rate for the confusion matrix is 0.00. This indicated that the categories were 100 percent accurate, 100 percent sensitive, 100 percent specific, and 100 percent precious. This performance measure was calculated using the testing data (132 samples) from the entire sample data. As previously said, any evaluation method may attain 100 percent if there is no error or misleading throughout the confusion matrix between actual and predicted quality.

There are several factors that cause the calculation of the performance measure to obtain a value of 100 percent. First, the implementation of detailed data preprocessing is required before using the dataset in a study. For this research, the dataset used has gone through several stages of data preprocessing using several techniques such as Pearson's correlation, natural logarithm, min. max. normalization, Principal Component Analysis, and many more [1]. Second, the determination of the appropriate use of intelligent techniques in constructing model classification also plays an important role. Each intelligent technique created has different capabilities. The same goes for the SVM technique. The OVO strategy is used to increase its capability from a two-class binary classification technique to a multiclass classification. As a result, the process of classifying four different grades of agarwood oil can be implemented successfully.

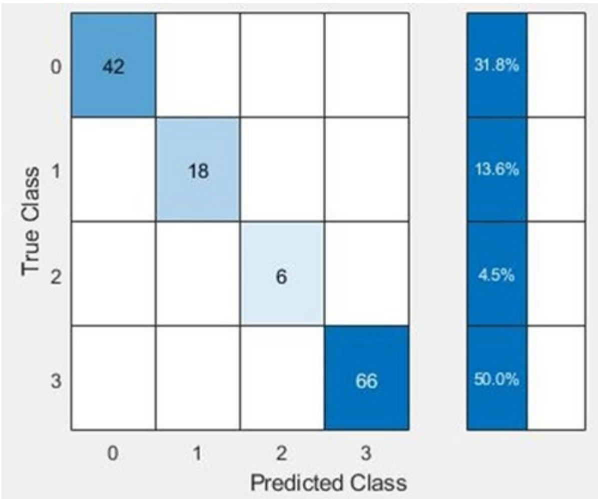

Fig. 11. Table analysis of confusion matrix.

IV. CONCLUSION

The Agarwood oil quality classification that is implemented now only focuses a lot on high and low quality classifications. Practically no one has given the classification of agarwood oil more than these two qualities. In this study, a classification of four different qualities was carried out, and the simulation was successful by using the appropriate intelligent technique. Basically, the use of the SVM technique to make a classification of more than two classes is impossible because it exceeds the limitations of this technique. The SVM technique is only capable of classifying two different classes at a time. This means, if only using the SVM technique, the datasets used in this study need to be divided manually into 6 different datasets and the classification process needs to run one by one at a time. But with the addition of the OVO strategy, the direct use of datasets having four different classes can be used. As a result, the OVO strategy is used in the creation of this model classification. This study has effectively shown the development of the modelling of agarwood oil classification using Multiclass Support Vector Machine based on the One Versus One (OVO) strategy. This developed model is very efficient, fast, easy, and, importantly, safe for anyone who handles it.As discussed in the results and discussion section, two important factors need to be focused on to implement the classification of agarwood oil qualities. First, focus needs to be given when data preprocessing is implemented. This is important to ensure that the abundance of the significant chemical compounds is clear to facilitate the classification of the model later to make comparisons between the qualities. Second, the use of intelligent techniques must be in line with the requirements required to classify the dataset used. Its capability must reach a higher level or at least be in line with the criteria required by the dataset used. Thus, the results will be evident when a performance measure is made against the model classification. This is evident when the calculation of the performance measure recorded for this study obtained a value of 100 percent for accuracy, sensitivity, specificity, and precison.The results of this study will certainly be able to provide benefits to studies that will be implemented in the future. To ensure a more accurate classification process, it is possible to study the properties of agarwood oil through the abundance of significant chemical compounds. Next, the data can be applied to this MSVM classification model to see whether the accuracy of its classification is the same or not. Also, the value of support vectors is predicted to be much more accurate than the currently used dataset.

ACKNOWLEDGMENT

This research is funded by Institute of Research Management and Innovation (IRMI), Universiti Teknologi MARA Shah Alam, Selangor and Universiti Teknologi MARA, Cawangan Johor, Kampus Pasir Gudang under Grant No: 600-IRMI/FRGS 5/3 (224/2019). The authors would like to express their gratitude to all of the staff from School of Electrical Engineering, College of Engineering, UiTM Shah Alam, Bioaromatic Research Centre of Excellence (BARCE), UMP and Natural Product Division, Forest Research Institute Malaysia (FRIM) for their contributions to this research at various stages.

REFERENCES

[1] M. H. Haron, "Agarwood Oil Quality Grading Model Using Self-Organizing Map (SOM)", PhD Dissertation, Faculty of Electrical Engineering, UniversitiTeknologi MARA (UiTM), 2020.

[2] S. R. Adhikari, K. Pokhrel, S. D. Baral, "Economic Value of Agarwood and Its Prospects of Cultivation", International Journal of Applied Sciences and Biotechnology (IJASBT), Vol. 9(1), pp. 23–31, 2021.

[3] N. Ismail, "ANN Modelling for Agarwood Oil Significant Compounds for Oil Quality Discrimination", PhD Dissertation, Faculty of Electrical Engineering, UniversitiTeknologi MARA (UiTM), 2015.

[4] N. S. Zubir, M. A. Abas, N. Ismail, N. A. M. Ali, M. H. F. Rahiman, N. K. Mun, M. N. Taib and S. N. Tajuddin, "Pattern Classifier of Chemical Compounds in Different Qualities of Agarwood Oil Parameter using Scale Conjugate Gradient Algorithm in MLP," 2017 IEEE 8th Control Syst. Grad. Res. Colloquium, ICSGRC 2017 - Proc., no. March, pp. 10–12, 2017.

[5] S. Wang, Z. Yu, C. Wang, C. Wu, P. Guo and J. Wei, "Chemical Constituents and Pharmacological Activity of Agarwood and Aquilaria Plants", MDPI, 2018.

[6] N. Ismail, H. Ariffin, M. H. F. Rahiman, M. N. Taib, N. A. M. Ali and S. N. Tajuddin," Statistical Learning BSVM Model to the Problem of Agarwood Oil Quality Categorization", 2018 IEEE Conference on Systems, Process and Control (ICSPC 2018), 14 – 15 December 2018, Melaka, Malaysia.

[7] D. T. T. Thuy, T. T. Tuyen, T. T. T. Thuy, P. T. H. Minh, Q. T. Tran, P. Q. Long, D. C. Nguyen, L. G. Bach, N. Q. Chien, "Isolation Process and Compound Identification of Agarwood Essential Oils from Aquilaria Crassna Cultivated at Three Different Locations in Vietnam", MDPI, 2019.

[8] X. Liu, B. F. Zhang, L. Yang, G. X. Chou, Z. T. Wang, "Whole-tree agarwood-inducing technique: an efficient novel technique for producing high-quality agarwood in cultivated Aquilaria sinensis tree", Molecules, vol. 18, pp. 3086-3106,2013.

[9] M. Antonopoulou, J. Compton, L. S. Perry, R. Al-Mubarak, "The trade and use agarwood (oudh) in the United Arab Emirates", Traffic Southeast Asia, Geneva: The CITES secretariat, 2010.

[10] M. J. Abdin, "The agarwood oil industry: yet to utilize in Bangladesh", Int. J. Econ. Manag. Sci, vol. 2, pp 163-266, 2014.

[11] S. Akter, M. T. Islam, M. Zulkefeli, S. I. Khan, "Agarwood production – a multidisciplinary field to be explored in Bangladesh", Int. J. Pharm. Life. Sci. vol. 2, pp. 22-32, 2013.

[12] N. Ismail, H. Ariffin, M. H. F. Rahiman, M. N. Taib, N. A. M. Ali and S. N. Tajuddin," Statistical Learning BSVM Model to the Problem of Agarwood Oil Quality Categorization", 2018 IEEE Conference on Systems, Process and Control (ICSPC 2018), 14 – 15 December 2018, Melaka, Malaysia.

[13] M. Aqib et al., "Agarwood Oil Quality Classification using Support Vector Classifier and Grid Search Cross Validation Hyperparameter Tuning," Int. J. Emerg. Trends Eng. Res., vol. 8, no. 6, pp. 6–11, 2020.

[14] N. A. B. Sidik, "Comparison of Gaharu (Aquilaria Malaccensis) Essential Oil Compasition Between Each Country", Bachelor of Chemical Engineering, Faculty of Chemical Engineering and Natural Resources, University Malaysia Pahang, 2008.

[15] Nor Azah M.A, Chang Y.S., Mailina J., Abu Said A, Abd. Majid J. SaidatulHusni S., Nor Hasnida H. and Nik Yasmin Y, "Comparison of Chemical Profiles of Selected Gaharu Oils from Peninsular Malaysia", The Malaysian Journal of Analytical Sciences, vol. 12, pp. 338 – 340, 2008.

[16] R. Naef, "The volatile and semi-volatile constituents of agarwood, the infected heartwood of Aquilaria species: a review", Flavour and Fragrance Journal, vol. 26, pp. 73-87, 2011.

[17] R. Kalra and N. Kaushik, "A review of chemistry, quality and analysis of infected agarwood tree (Aquilaria sp.)," Phytochem. Rev., vol. 16, no. 5, pp. 1045–1079, 2017.

[18] K. A. A. Kamarulzaini, N. Ismail, M. H. F. Rahiman, M. N. Taib, N. A. M. Ali and S. N. Tajuddin," Evaluation of RBF and MLP in SVM Kernel Tuned Parameters for Agarwood Oil Quality Classification", 2018 IEEE 14th International Colloquium on Signal Processing & its Application (CSPA 2018), 9 -10 March 2018, Penang, Malaysia.

[19] M. A. N. Azah, S. S. Husni, J. Mailina, L. Sahrim, J. Abdul Majid, and Z. Mohd Faridz, "Classification of agarwood (gaharu) by resin content," J. Trop. For. Sci., vol. 25, no. 2, pp. 213–219, 2013

[20] M. A. A. Ngadilan, N, Ismail, M. H. F. Rahiman, N. A. M. Ali and S. N. Tajuddin," Radial Basis Function (RBF) tuned Kernel Parameter of Agarwood Oil Compound for Quality Classification using Support Vector Machine (SVM)", 2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC 2018), 3 – 4 August 2018, Shah Alam, Malaysia.

[21] H. Jantan, I. M. Yassin, A. Zabidi, and N. Ismail, "Differentiation of Agarwood Oil Quality Using Support Vector Machine (SVM)," J. Eng. Appl. Sci., vol. 12 (15), no. 3810–3812, 2017.

[22] M. A. A. Aziz, N. Ismail, I. M. Yassin, A. Zabidi, and M. S. A. M. Ali, "Agarwood oil quality classification using cascade-forward neural network," Proc. - 2015 6th IEEE Control Syst. Grad. Res. Colloquium, ICSGRC 2015, pp. 112–115, 2016.

[23] N. Coskun and T. Yildirim, "The effects of training algorithms in MLP network on image classification," Proc. Int. Jt. Conf. Neural Networks, vol. 2, no. see 17, pp. 1223–1226, 2003.

[24] M. M. Islam, J. M. Kim, "Reliable multiple combined fault diagnosis of bearings using heterogeneous feature models and multiclass support vector machines", Reliability Engineering System Safety, vol. 184, pp 55-66, 2019.

[25] M. Sun, K. Liu, Q. Hong, "An ECOC Approach for Microarray Data Classification Based on Minimizing Feature RelatedComplexities", 2017 10th International Symposium on Computational Intelligence and Design (ISCID), vol. 1, 2017.

[26] Z. Tianyun, W. Jiajun, L. Kunhong, W. Beizhan, H. Qingqi, "Multiclass Microarray Data Classification Based on SA-ECOC", 2017 10th International Symposium on Computational Intelligence and Design (ISCID), vol. 1, 2017.