International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V69I6P218

Volume 69 Issue 6, 121-125, June 2021

One Versus All Strategies of Multiclass SVM in Modeling Agarwood Oil Quality Classification

Aqib Fawwaz Mohd Amidon[1], Noratikah Zawani Mahabob[1], Nurlaila Ismail[1#], Zakiah Mohd Yusoff[2], Mohd Nasir Taib[3]

[1]*School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia*
[2]*School of Electrical Engineering College of Engineering, Universiti Teknologi MARA, Cawangan Johor, Kampus Pasir Gudang, Malaysia*
[3]*Malaysia Institute of Transport (MITRANS), Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia*

[#]nurlaila0583@uitm.edu.my

Abstract *— Agarwood oil is one of the most beneficial oil to the world community with a high demand. It is beneficial due to the variety of usages such as incense, traditional medicine, and perfumes. However, there has been a lack of research on the development of agarwood oil because there is no any standard grading model of agarwood oil was implemented. As a solution forms, it is very important to come out with a standard of quality classification model for agarwood oil grading's. By continuing of the research for the development of this standard, specific algorithm function has been used to make sure the ability of this model is totally not in doubt. Support vector machine (SVM) has been chosen as a main model and for the specific function algorithm that has been chosen was multiclass function. Then, in the function, the one versus all (OVA) strategies has been used to make multiclass work and can be applied on SVM. The analysis work has involving the data taken from the previous researcher that consists of four classes of agarwood oil quality's samples which are low, medium low, medium high and high quality. So, the output was the classification of quality between low, medium low, medium high or high quality while the input was the abundances (%) of compounds. The desk research has been conducted by using MATLAB software version r2020a for the simulation platform. The result showed that the model by using multiclass function has pass the performance criteria standard. The verdict in this research for sure will be valuable for the future research works of agarwood oil areas, especially quality classification part.*

Keywords *— Agarwood oil, Mutliclass, One Versus All, Support Vector Machine, Performance criteria.*

I. INTRODUCTION

Agarwood or known as chenxiang in Chinese and called agalloch, aloeswood, eaglewood, kanankoh, jinkoh, kalambak or gaharu in different regions, is a non-timber forest product with very highly prized for its values as herbal medicine, aromatic material, ceremonies and many more, depend on the different cultures and religious [1]-[6]. Agarwood oil, the fragrant resinous heartwood is a well-known essential oil originating from Aqualaria and several other plant species of the Thymelaeaceae family [1],[7]. The Aquilaria species were founded around 31 species around the worldwide in Malaysia, Indonesia, Thailand, Vietnam, Laos, Singapore, Philippines, Cambodia, Bangladesh, Papua New Guinea, China, Japan, and India [2],[4]-[6],[8]-[9].

The way of grading method is different according to countries [5],[10]. Some countries like Malaysia, Japan and India graded agarwood using two types of grading techniques. Malaysian researchers classify agarwood by naming them as kalambak and gaharu as well as in Japan grade agarwood by use kanankoh and jinkoh for high and low quality. However, India, they are using the alphabet from A to D or numbering from 1 to 4 according to color and infection level for the agarwood grading process [11]. The price, usability and others are fully depending on the quality of the agarwood oil either the quality is low, medium low, medium high or high [5]. In other word, the higher the quality of agarwood oil, the higher the price of the agarwood oil [5]. A review from other researchers' study [9],[12],[13], it was found that grading agarwood oil process has been conducted by only using human sensory panels. But this method can spark doubt because it is very incompetent, caused tiredness, and take time as well.

With the availability of data analysis technology nowadays, there is a platform where agarwood oil quality classification can be conduct only use their chemical profiles so that essential oil can be classified according to their respective classes (low, medium low, medium high or high quality) and the accurate result can be measured. Today, researchers use a modern technique for classifying nonlinear data including agarwood oil such as by using machine

learning including Support Vector Machine (SVM), Multilayer Perceptron (MLP), Artificial Neural Network (ANN), Linear Regression and various analysis techniques such as statistical analysis [14]-[17].

According to the techniques mentioned above, in this paper has focus on the Support Vector Machine as the main model to classification the grading of agarwood oil into low, medium low, medium high and high quality as recommended by [18]-[20].

II. METHODOLOGY

In this research, there are 660 of sample of data were used throughout the research [1]. These data samples have four different classes which include 210 samples of data for low quality, 90 samples of data for medium low quality, 30 samples of medium high quality and 330 samples of high quality. The data consists of eleven inputs which indicate the eleven selected significant compounds knows as C1=α-guaiene, C2=β-agarofuran, C3=ar-curcumene, C4=β-dihydro agarofuran, C5=ϒ-cadinene, C6=α-agarofuran, C7=10-epi-ɤ-eudesmol, C8=ɤ-eudesmol, C9=allo aromadendrene epoxide, C10=valerianol, C11=dihydrocollumellarin.

A. Experimental Set – up

All these data have been through data pre-processing stage which are normalizing data, randomizing the arrangement of data, and lastly separating into training data and testing data by using 80%:20% ratio as recommended by previous researchers [1] as mentioned on flowchart in Fig. 1. The simulation has used Holdout partition to separated data as recommended.

After that, Support Vector Machine (SVM) model have been developed by using standard template that have been designed for nonlinear data. The standard template using SVM as method, classification as the type of analysis, and 'gaussian' as kernel function parameter by having the ability to analyze nonlinear data [20],[21]. Then, to make sure the classification successful, multiclass classification method based on binary classifiers have been used [23]-[26]. In the multiclass classification, the One versus All (OVA) strategies has been chosen as multiclass method of machine learning (A. Rocha & S. K. Goldenstein 2014). With that, each of the quality will be classify by comparing the quality and other qualities as Low vs All (LvsA), Medium Low vs All (MLvsA), Medium High vs All (MHvsA) and High vs All (HvsA).

Next, all 528 of training data have been involve during develop the SVM model above. But 132 of testing data have been hold first for testing the model and analyze the model by following the performance criteria as standard pass evaluation for SVM model. The result of testing has been put in result and discussion part.

The simulation of intelligent model has been conducted by using Matlab software version r2020a to classify four different quality of agarwood oil.

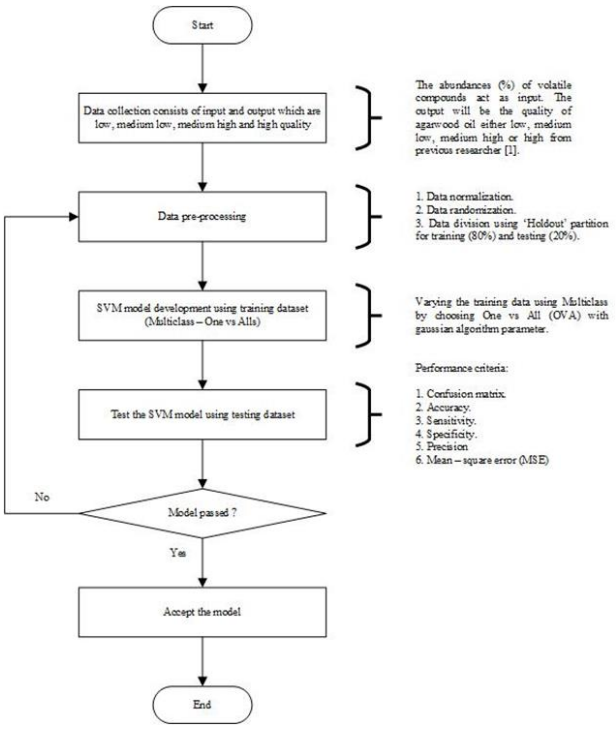


Fig. 1 Detail of experimental set – up

III. RESULT AND DISCUSSION

As mentioned in materials and methods part, there were 660 samples of agarwood oil compound data used in the analysis work. The sample of data consists of four different quality which are low, medium low, medium high and high quality. The data samples then were analyzed and constructed into SVM model. Next, One versus All (OVA) of multiclass function have been applied on the SVM model to classify the quality of data.

A. Generate Support Vectors

The outcome from the applying the OVA have formed different position of support vectors as shows in Fig. 2 – 5 according to the type of qualities.

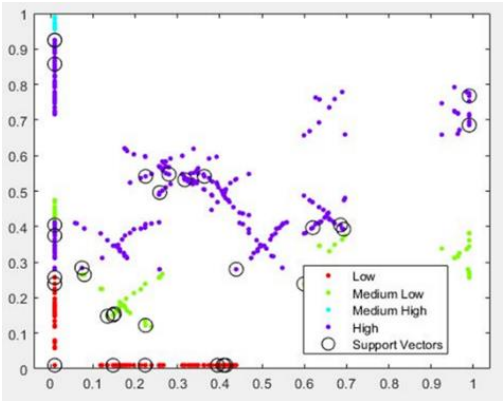


Fig. 2 The output of support vectors for (LvsA)

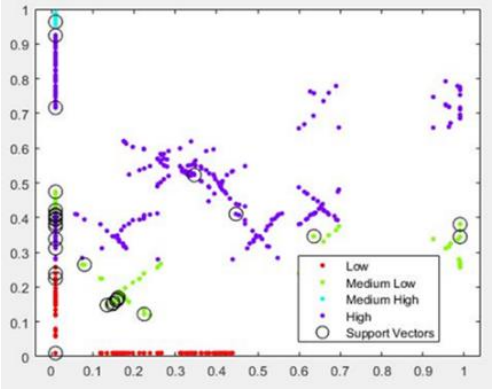


Fig 3. The output of support vectors for (MLvsA)

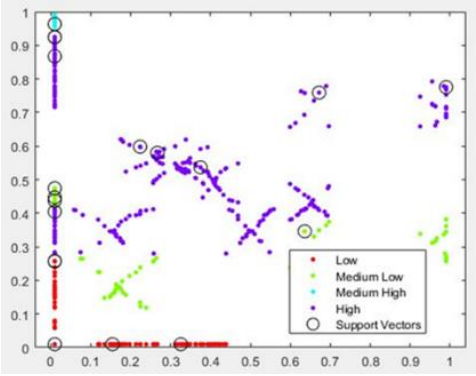


Fig 4. The output of support vectors for (MHvsA)

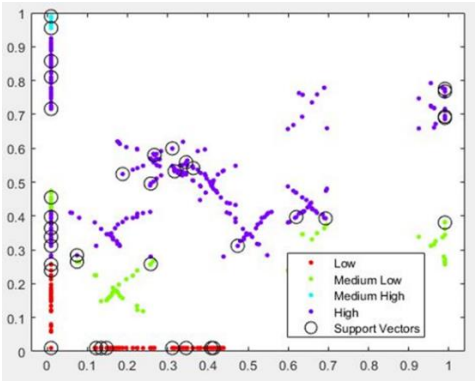


Fig 5. The output of support vectors for (HvsA)

Based on the Table I, the table showed the minimum and maximum value of support vectors according to the quality of agarwood oil based on the 11-chemical compound. For α-guaiene, the minimum and maximum value of support vectors per quality are Low vs All have 0.01 as minimum value and 0.98 as maximum value, Medium Low vs All have 0.01 as minimum value and 0.99 as maximum values which is same value to Medium High vs All and High vs All. For β-agarofuran, Low vs All have 0.01 as minimum value and 0.93 as maximum value, Medium Low vs All have 0.01 as minimum value but 0.99 as maximum value which is same value to Medium High vs All and High vs All. For ar-curcumene, Low vs All have 0.01 as minimum value and 0.99 as maximum value which is same value to Medium High vs All and High vs All but Medium Low vs All have 0.01 as minimum value and 0.82 as maximum value. For β-dihydro agarofuran, all qualities have 0.01 as minimum value but for maximum value, only Medium Low vs All have 0.58 as maximum value, the other quality have 0.99 as maximum value. For ϒ-cadinene, all qualities have 0.01 as minimum value but for maximum value, every quality have different values. Low vs All is 0.97, Medium Low vs All is 0.83, Medium High vs All is 0.92 and High vs All is 0.97. For α-agarofuran, all qualities have 0.01 as minimum value, but every quality have different value for maximum, Low vs All is 0.97, Medium Low vs All is 0.83, Medium High vs All is 0.92 and High vs All is 0.97. Next, for 10-epi-ɤ-eudesmol, Low vs All have 0.31 as minimum value and 0.81 as maximum value., Medium Low vs All have 0.28 as minimum value and the maximum value is 0.92, Medium High vs All have 0.35 as minimum value and 0.97 as maximum value and for High vs All, it has 0.01 as minimum value and 0.99 as maximum value. For ɤ-eudesmol, all qualities have 0.01 as minimum value but each of the quality have different value for maximum value. Low vs All is 0.85, Medium Low vs All is 0.96, Medium High vs All is 0.90 and High vs All is 0.99. For allo aromadendrene epoxide, all qualities have 0.01 as minimum value but not for for maximum value. Low vs All is 0.99, Medium Low vs All is 0.80, Medium High vs All is 0.91 and High vs All is 0.99. For valerianol, all qualities have 0.01 as minimum value but not for for maximum value. Low vs All is 0.97, Medium Low vs All is 0.67, Medium High vs All is 0.66 and High vs All is 0.97. Lastly, for dihydrocollumellarin, all qualities have 0.01 as minimum value but not for for maximum value. Low vs All is 0.99, Medium Low vs All is 0.52, Medium High vs All is 0.96 and High vs All is 0.99.

TABLE I

MINIMUM AND MAXIMUM VALUE OF SUPPORT VECTORS

	Chemical comp 1		Chemical comp 2		Chemical comp 3		Chemical comp 4		Chemical comp 5		Chemical comp 6	
	α-guaiene		β-agarofuran		ar-curcumene		β-dihydro agarofuran		ϒ-cadinene		α-agarofuran	
	Min. value	Max. value	Min. value	Max. value	Min. value	Max. value	Min. value	Max. value	Min. value	Max. value	Min. value	Max. value
(LvsA)	0.01	0.98	0.01	0.93	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.97
(MLvsA)	0.01	0.99	0.01	0.99	0.01	0.82	0.01	0.58	0.01	0.70	0.01	0.83
(MHvsA)	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.92
(HvsA)	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.97

	Chemical comp 7		Chemical comp 8		Chemical comp 9		Chemical comp 10		Chemical comp 11	
	10-epi-ɤ-eudesmol		ɤ-eudesmol		allo aromadendrene epoxide		valerianol		dihydrocollumellarin	
	Min. value	Max. value	Min. value	Max. value	Min. value	Max. value	Min. value	Max. value	Min. value	Max. value
(LvsA)	0.31	0.81	0.01	0.85	0.01	0.99	0.01	0.97	0.01	0.99
(MLvsA)	0.28	0.92	0.01	0.96	0.01	0.80	0.01	0.67	0.01	0.52
(MHvsA)	0.35	0.97	0.01	0.90	0.01	0.91	0.01	0.66	0.01	0.96
(HvsA)	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.97	0.01	0.99

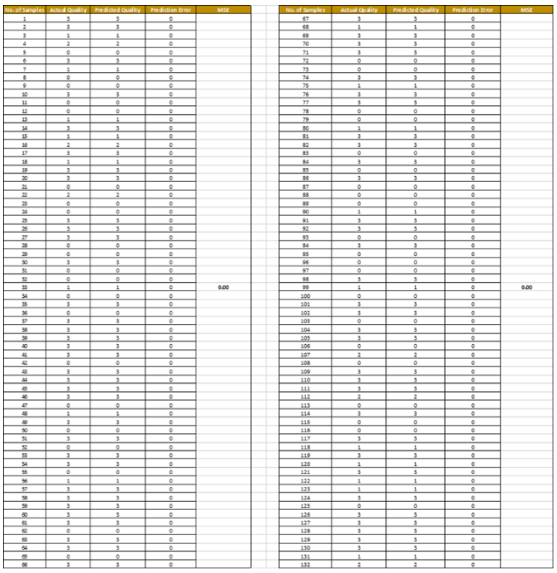
B. Performance Criteria

Table 2 provides the confusion matrix of the SVM model between actual quality and predicted quality for four different qualities of agarwood oil by using testing data. It shows that, there is no any misleading during predicted and actual quality. With that, error rate for confusion matrix is 0.00 and for sure Mean- square error also have 0.00 value as

show in Table 2. Moreover, on confusion matrix table based on Fig. 6 have been proven there is no misleading happen between predicted and actual quality.

TABLE II

ACTUAL QUALITY, PREDICTED QUALITY, PREDICTED ERROR, AND MEAN-SQUARE ERROR USING TESTING DATA



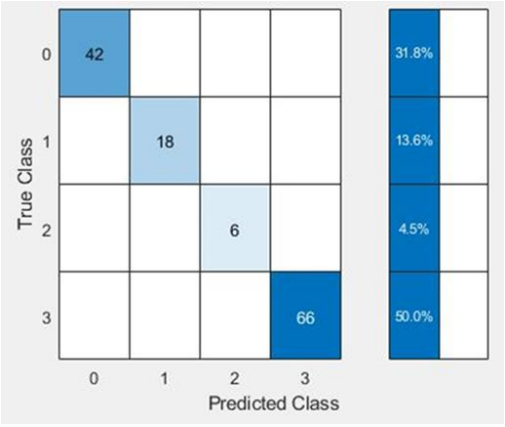


Fig. 6 Confusion matrix table

Table 3 shows that each of the classes have 100% of accuracy, 100% of sensitivity, 100% of specificity and 100% for precious. This performance criteria have been measure based on the testing data which is 20% (132 samples) from total of sample of data. As researcher mention before, each of performance criteria can achieve 100% when there is no any error or misleading during confusion matrix between actual quality and predicted quality.

TABLE III

PERFORMANCE CRITERIA

	LvsA	MLvsA	MHvsA	HvsA	Avg Acc	Avg Sen	Avg Spec	Avg Pre
Accuracy	100%	100%	100%	100%	100%			
Sensitivity	100%	100%	100%	100%		100%		
Specificity	100%	100%	100%	100%			100%	
Precious	100%	100%	100%	100%				100%

IV. CONCLUSIONS

Various studies have made an effort to show that the insufficiency of classification of quality using traditional method can affect human health. By developing of the modelling of agarwood oil classification using Support Vector Machine (SVM) has successfully presented on this paper. During SVM process by support with multiclass classification method, the classification had been run using hyperplane separating in SVM for the result. From the data, accuracy yielded to 100% which are very significant and also reach 100% of sensitivity. With that, 0.00 of error rate for the model result have been achieved. On the other hand, the SVM model in this study keep strongly proven the ability on handling the variation of input data created from the abundances of eleven significant chemical compounds and separate the nonlinearly data into training and testing data. then, measure the output of data by classify the data into four classes of quality, low, medium low, medium high, and high quality.

The outcome of this study benefits other analysts on evaluating the oil quality especially agarwood oil. By developing this intelligent model, SVM and implementing one versus all of multiclass function will be significant for future study especially in agarwood oil classification. In future work, it is prescribed to design the agarwood oil model to classify more than two classes of agarwood oil quality. Their result afterward can be compared and assessed.

ACKNOWLEDGMENT

This research is funded by Ministry of High Education (MOHE), Malaysia under FRGS Grant No: 600-IRMI/FRGS 5/3 (031/2019). The authors would like to thank Dr Mohd Hushnie Haron, Dr Nor Azah Mohd Ali, Assoc. Prof. Dr. Saiful Nizal Tajuddin and all the staff involved from the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA (UITM) Shah Alam, Natural Product Division, Forest Research Institute Malaysia (FRIM) and Bioaromatic Research Centre of Excellence (BARCE), Universiti Malaysia Pahang (UMP) for their support at various stages for this study.

REFERENCES

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

[2] N. Ismail, ANN Modelling for Agarwood Oil Significant Compounds for Oil Quality Discrimination, PhD Dissertation, Faculty of Electrical Engineering, Universiti Teknologi Mara (UiTM), (2015).

[3] 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, (2017) 10–12.

[4] 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).

[5] 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.

[6] 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).

[7] M. H. Haron, M. N. Taib, N. Ismail, N. A. M. Ali, and S. N. Tajuddin, Statical Analysis of Agarwood Oil Compounds based on GC-MS Data, (2018) 9th IEEE Control and System Graduate Research Colloquium (ICSGRC 2018), 3 – 4 August 2018, Shah Alam, Malaysia.

[8] N. Ismail, N. A. M. Ali, M. Jamil, M. H. F. Rahiman, S. N. Tajuddin, and M. N. Taib, A review study of agarwood oil and its quality analysis, J. Teknol. (Sciences Eng., 68 (1) (2014) 37–42.

[9] 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.

[10] 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., 8(6) (2020) 6–11.

[11] R. Kalra and N. Kaushik, A review of chemistry, quality and analysis of infected agarwood tree (Aquilaria sp.), Phytochem. Rev., 16 (5) (2017) 1045–1079.

[12] 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.,. 25 (2) (2013) 213–219.

[13] 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.

[14] H. Jantan, I. M. Yassin, A. Zabidi, and N. Ismail, Differentiation of Agarwood Oil Quality Using Support Vector Machine (SVM), J. Eng. Appl. Sci., 12 (15) (2017) 3810–3812.

[15] 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, (2016) 112–115.

[16] H. Taud and J. Mas, Multilayer Perceptron (MLP), in Geomatic Approaches for Modeling Land Change Scenarios, ed: Springer, (2018) 451-455.

[17] R. Kamminana, V. Kambagowni, Modeling and Optimization of Process Parameters of Friction Stir Welding of Al-Li Alloy AA2050 by Response Surface Methodology, International Journal of Engineering Trends and Technology 69(5) (2021) 208-227.

[18] H. Kerdegari, K. Samsudin, A. R. Ramli, and S. Mokaram, Evaluation of fall detection classification approaches, in Intelligent and Advanced Systems (ICIAS), 2012 4th International Conference (2012) 131-136.

[19] N. Ismail, M. H. F. Rahiman, M. N. Taib, M. Ibrahim, S. Zareen, and S. N. Tajuddin, A review on agarwood and its quality determination, Proc. - 2015 6th IEEE Control Syst. Grad. Res. Colloquium, ICSGRC 2015. (2016) 103–108.

[20] 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 13th International Colloquium on Signal Processing * its Applications (CSPA 2017), (2017) 10-12 Penang, Malaysia.

[21] Y. Zhang and R. Li, Selection of the Parameter in Gaussian Kernels in Support Vector Machine, 2017th 2nd IEEE International Conference on Cloud Computing and Big Data Analysis, (2017).

[22] N. S. Zubir, N. Ismail, M. H. F. Rahiman, M. N. Taib, N. A. M. Ali and S. N. Tajuddin, Polynomial tuned Kernel Parameter in SVM of Agarwood Oil for Quality Classification, 2018 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS 2018), (2018), Shah Alam, Malaysia.

[23] A. Mathur and G. M. Foody, Multiclass and Binary SVM Classification: Implications for Training and Classification Users, IEEE Geoscience and Remote Sensing Letters, 5 (2) (2008).

[24] A. Rocha and S. K. Goldenstein, Multiclass from Binary: Expanding One-Versus-All, One-Versus-One and ECOC-Based Approaches, IEEE Transactions on Neural Networks and Learning Systems, 25 (2) 2014.

[25] Irina Klyueva, Improving Quality of the Multiclass SVM Classification Based on the Feature Engineering, 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency (SUMMA), (2019).

[26] S. Daengduang and P. Vateekul, Applying One-Versus-One SVMs to Classify MutlciLabel Data with Large Labels Using Spark, 2017 9th International Conference on Knowledge and Smart Technology (KST), (2017).