ISSN 2347 - 3983
Volume 8. No. 9, September 2020
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter93892020.pdf
https://doi.org/10.30534/ijeter/2020/93892020

SVM Modelling of Agarwood Oil Quality Grading using Radial Basis Function (RBF) and Sequential Minimal Optimization (SMO) Learning Algorithms

Muhammad Haziq Haikal Hisham[1], Nurlaila Ismail[1*], Nor Azah Mohd Ali[2], Saiful Nizam Tajuddin[3], Mohd Nasir Taib[1,4]
[1]Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia.
*nurlaila0583@uitm.edu.my
[2]Herbal Product Development, Natural Product Division, Forest Research Institute Malaysia, Malaysia.
norazah@frim.gov.my
[3]Bioaromatic Research Centre of Excellence, Universiti Malaysia Pahang, Gambang, Pahang,
Malaysia.saifulnizam@ump.edu.my
[4]Malaysia Institute of Transport (MITRANS), Universiti Teknologi MARA, Selangor, Malaysia.
dr.nasir@uitm.edu.my

ABSTRACT

Agarwood Oil also known as Gaharu Oil is an expensive oil with extreme demand in the world trading especially in Japan, China and the Middle East countries. Currently, the grading of the agarwood oil quality only can be done by trained human graders by physically checked the color, odour and the resin content. However, this technique is limited due to human body limitation and not too accurate as it is supposed to be. To improve the problem faced by the existing method, the grading technique using Sequential Minimal Optimization (SMO) and Radial Basis Function (RBF) in Support Vector Machine (SVM) was conducted. The works involved of data collection, data pre-processing, SVM model development and testing of the developed SVM model. The finding showed that both RBF and SMO successfully can grade the agarwood oil quality due to their accuracies is above than 80 % and error rate or MSE is close to 0. Thus, the technique presented in this paper proved its capability in overcoming the constraint of human trained grader and benefit as well as contribute to the agarwood oil industry especially its oil quality grading.

Key words: SMO, RBF, SVM, Oil quality, Agarwood oil, Classification

INTRODUCTION

Agarwood, also known in Malaysia as Gaharu, is a fragrant and valuable foreign plant harvested from Aquilaria and Gyrinops tree species [1]. Agarwood oil only can be obtained from the Aquilaria trees that are infected by a certain mold or attacked by insects and bacteria. This resinous heartwood of a tree will differentiate between an infected tree with non-infected tree. Four species of agarwood which are: i.e. A. filaria (Oken) Merr. A. sinensis (Lour.) Spreng, A. crassna Pierre ex Lecomte and A. malaccensis Lam. [2].

Back in 5th century, agarwood was used a lot in Chinese medicines, which are about beyond 1500 variation of preparations of Chinese medical stuffs. The bitterness taste of agarwood is used as to cure gastric issues, sedative and carminative, high fever, rheumatism and coughs. Not limited to that, agarwood has been used as incense in Islamic, Hindu and Buddhist ceremonies for centuries [2].

In agarwood trading field, estimated that beyond 18 countries throughout Middle East and Southeast Asia have involved. Although every country has their own medical systems that has various agarwood grades, some indicators like sinkage, scent, color and resin have been evaluated in evaluating agarwood worldwide. Many countries use agarwood oil wit term quality; high and low and some countries use agarwood oil with term grade; A, B, C and D [3]. Each of agarwood parts has their own preference by traders such as, the resin content and chemical compounds are important for medical uses, aroma and color are preferred in religion purpose, and aroma and shape are important in collection purpose [4].

As agarwood oil is widely used and getting high demand in market, thus various methods have been discovered to determine the agarwood oil quality. Most of them are based on physical properties of the oil and seldom of them are based on chemical properties of the oil. Therefore, this study proposed another method which will utilize the chemical properties of the oil. The proposed technique also implemented the machine learning in classifying the agarwood oil quality. Which is Support Vector Machine (SVM) [5].

SVM created a separating hyperplane of agarwood oil compounds in a hyperplane region. The SVM is a computer algorithm that learns to set labels for objects with an example [5]. Vladimir Vapnik together with his team invented and first presented at the Computational Learning Theory (COLT) 1992 conference with the paper. In 1995, Vapnik defined Support Vector Machines (SVM) as one of machine learnings that distinguish all data points in one class from those in other class, by finding the right hyperplane [6]

Radial Basis Function (RBF) is a kernel function that been used in classification method. It is a network architecture that uses a radial basis feature as neuron transmission or activation [9-13]. There are many RBF but normally a Gaussian function is used. The RBF is a feedforward network which can be used in places where a conventional network of mixtures can be found. This includes until the final pattern classification from function estimation [14].

The main point of using RBF is to get the nodes with the minimum number to prevent wastage of effort in lengthy calculations. In recent years, major evolutions have been seen for RBF method, which is a technique for interpolation in a high dimensional space [15].

Sequential Minimal Optimization (SMO) is popular training algorithm and a simple way that can easily solve the SVM quadratic programming (QP) tasks without using numerical QP optimization and without any extra matrix storage at all [18]. It is different from other methods, in every step, it opts to solve the least optimization problem possible. Normally, typical SVM QP problem includes two Lagrange multipliers. By choosing two different Lagrange multipliers to jointly optimize at every step and get the best values for the multipliers, SVM will be updated by SMO to check the best rates.

In conjunction to that, the study presented the grading technique of agarwood oil quality with the uses of RBF and SMO learning algorithms during the development of the SVM modelling.

2. METHOD

The collaboration between the Forest Research Institute Malaysia (FRIM) and Bio Aromatic Research Centre of Excellence (BARCE) of University of Malaysia Pahang (UMP) has provided the agarwood oil sample for the project by the previous researcher [20]. By using RBF and SMO in SVM modelling, the quality of the samples evaluated and graded according to their compound quality.

Figure 1 shows the flow chart of the SVM model development carried out in this study. The whole process started by getting all the raw data from the previous researcher [20]. Then, data pre-processing had been carried out, which means the data are normalized, randomized and split into training and testing datasets. The next process is where the SVM model were developed using the training dataset which are 80% of all data. After that, 20% of all data are the testing dataset that were used to test the built network. It will only be accepted if the model passed, however if not, it will undergo data pre-processing again. To pass, the model needs to fit the performance criteria such as specificity, precision, sensitivity, confusion matrix and accuracy. Usually, the built model that achieved above than 80% of accuracy [20] can be accepted and declared as reliable to differentiate the agarwood oil quality.

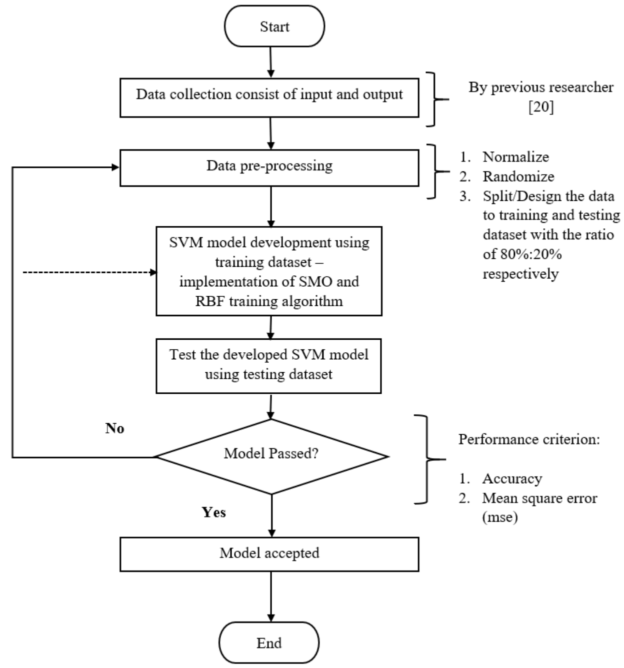

Figure 1: Flowchart of the SVM model development

3. RESULTS AND DISCUSSION

The whole simulation was conducted using MATLAB r2018a. The code started with the load of the data into the Matlab workspace. Then, followed by the set of the initial weights, declared the input as ‘xdata’ and output as ‘group’ as well as its classification error. Table I and Table II shows the support vectors for seven compounds conducted using Sequential Minimal Optimization (SMO) and Radial Basis Function (RBF), respectively. In Table 1, 13 support vectors were produced by using SMO algorithm. Their values are varied to each other from C1 to C7 (C1 is β-agarofuran, C2 is α-agarofuran, C3 is 10-epi-ϒeudesmol, C4 is ϒ-eudesmol, C5 is longifolol, C6 is hexadecanol and C7 is eudesmol).Firstly, the compound C1 and C2, the minimum vector of 0.00 were found the same at vector 1, 2 and 4. The maximum vector of 0.4200 at vector 13 for C1 and the maximum vector of 0.6100 was found at vector 11 for C2. For compound C3, the minimum vector of 0.4900 was found at vector 11 and 10.2100 is the maximum vector found at vector 2. Next, compound 4 has the minimum vector of 0.00 at vector 1 and vector 2, while the maximum vector of 2.9000 was found at vector 11. Subsequently, the minimum vector of 0.0000 were found at vector 2 and vector 13, 0.1700 is the maximum vector found at vector 1 for compound C5. Compound C6 showed the minimum vector of 0.0000 at vector 1, 2 and 8, while the maximum vector of 0.1600 was found at vector 6. Lastly, for compound C7, the minimum vector of 0.0000 were found at vector 1 and vector 2 and 3.0500 was found as the maximum vector at vector 11.

Table 1: Support vectors for seven compounds using Sequential Minimal Optimization (SMO)

No	C1	C2	C3	C4	C5	C6	C7
1	0.0000	0.0000	6.4200	0.0000	0.1700	0.0000	0.0000
2	0.0000	0.0000	10.2100	0.0000	0.0000	0.0000	0.0000
3	0.0500	0.0600	6.4000	0.0100	0.1500	0.0100	0.1100
4	0.0000	0.0000	10.1100	0.0400	0.0700	0.1400	0.1500
5	0.3500	0.0400	1.2600	1.6200	0.0200	0.0500	0.1000
6	0.1500	0.0700	6.5100	0.0600	0.1000	0.1600	0.0600
7	0.0300	0.0100	10.2000	0.0200	0.0200	0.0900	0.1100
8	0.3800	0.0100	1.2200	1.6100	0.0500	0.0000	0.0100
9	0.0800	0.0200	6.4000	0.0400	0.1400	0.1100	0.0100
10	0.0300	0.0400	6.4900	0.0300	0.0100	0.0600	0.0100
11	0.0900	0.6100	0.4900	2.9000	0.0300	0.0200	3.0500
12	0.0700	0.0100	6.5000	0.0700	0.1600	0.0300	0.0700
13	0.4200	0.0200	1.2400	1.6200	0.0000	0.1000	0.0500

Refer to Table 2, 33 support vectors were produced by using RBF algorithm. Their values are varied to each other from C1 to C7 (C1 is β-agarofuran, C2 is α-agarofuran, C3 is 10-epi-ϒ-eudesmol, C4 is ϒ-eudesmol, C5 is longifolol, C6 is hexadecanol and C7 is eudesmol). Firstly, compound C1, the minimum vector of -0.7251 were found at vector 1, 3, 7 and 31, while the maximum vector is 2.984. Next, compound C2, the minimum vector of -0.7530 were found same at vector 1, 3, 4 and 7, while the maximum value of 2.3903 was found at vector 15. Then, the minimum vector of -1.1986 was found at vector 18 and the maximum vector of 2.8137 was found at vector 25 of the compound C3. For compound C4, at vector 1 and 3, the minimum vector of -1.3727 were found and the maximum vector of 1.4972 was found at vector 29. While for compound C5, the minimum vector of -0.2773 were found same at vector 2, 3, 4, 6, 19, 26 and 32, and the maximum vector of 3.8796 was found at vector 18. Besides, for compound C6, the minimum vector of -0.3605 were found same at vector 1, 2, 3, 4 and 16, while the maximum vector of 4.1199 was found at vector 10. Lastly, the minimum vector of -0.3564 were found same at vector 1, 3, 4 and 20, and the maximum vector of 3.8344 was found at vector 22.

Table 2: Support vectors for seven compounds using Radial Basis Function (RBF)

No	C1	C2	C3	C4	C5	C6	C7
1	-0.7251	-0.7530	-0.0023	-1.3727	-0.2346	-0.3605	-0.3564
2	-0.6718	-0.1654	-1.1265	-0.7991	-0.2773	-0.3605	0.8455
3	-0.7251	-0.7530	0.7173	-1.3727	-0.2773	-0.3605	-0.3564
4	-0.5334	-0.7530	-0.9840	-1.0408	-0.2773	-0.3605	-0.3564
5	-0.6985	-0.6942	-0.0061	-1.3707	-0.2396	-0.3474	-0.3123
6	-0.6825	-0.2242	-1.1359	-0.8032	-0.2773	-0.3211	0.8415
7	-0.7251	-0.7530	0.6983	-1.3645	-0.2597	-0.1766	-0.2963
8	-0.5387	-0.7138	-0.9821	-1.0408	-0.2723	-0.2948	-0.3163
9	-0.6452	-0.6844	0.0148	-1.3604	-0.2522	-0.1503	-0.3324
10	-0.4588	-0.6159	-1.1530	-0.7397	3.8218	4.1199	3.5499
11	-0.6133	-0.1557	-1.1265	-0.7950	-0.2748	-0.2028	0.8816
12	-0.7091	-0.7432	0.7154	-1.3686	-0.2723	-0.2422	-0.3123
13	0.4890	0.6865	-0.3289	1.0916	-0.2522	-0.2685	-0.2643
14	1.0109	1.8223	0.5217	0.8847	-0.2245	-0.2817	-0.2963
15	1.3943	2.3903	2.7719	-1.1535	-0.2622	-0.1503	-0.2803
16	-0.5227	-0.7432	-0.9897	-1.0429	-0.2647	-0.3605	-0.3524
17	-0.6825	-0.7334	-0.0061	-1.3645	-0.2421	-0.2160	-0.3524
18	-0.5014	-0.6453	-1.1986	-0.6824	3.8796	3.5812	3.6822
19	-0.6452	-0.2340	-1.1283	-0.8217	-0.2773	-0.3211	0.8415
20	-0.5227	-0.7040	-0.9878	-1.0306	-0.2622	-0.3342	-0.3564
21	-0.7091	-0.7138	0.0110	-1.3666	-0.2748	-0.2817	-0.3524
22	-0.7038	-0.7040	-1.1948	-0.7295	3.8294	3.8308	3.8344
23	2.9864	-0.7236	-0.5416	-0.1764	-0.2723	-0.2554	-0.2963
24	-0.6505	-0.5963	0.7078	-1.3563	-0.2295	-0.1766	-0.3444
25	1.3943	2.3805	2.8137	-1.0880	-0.2069	0.1782	-0.2683
26	-0.5334	-0.7432	-0.9878	-1.0286	-0.2773	-0.3342	-0.3524
27	-0.6878	-0.7432	0.0129	-1.3584	-0.2371	-0.3211	-0.3284
28	-0.7144	-0.1752	-1.1037	-0.7991	-0.2698	-0.3342	0.8375
29	-0.6985	-0.5767	-0.5568	1.4972	-0.2321	-0.2028	-0.2322
30	-0.7198	-0.5180	-0.5720	0.2702	-0.2497	-0.2948	-0.3444
31	-0.7251	-0.6061	0.7097	-1.3399	-0.2748	-0.1766	-0.3324
32	-0.5014	-0.7334	-0.9859	-1.0408	-0.2773	-0.2291	-0.3364
33	-0.2991	1.2054	0.9053	-0.4775	-0.2471	-0.0977	-0.2442

The performance of RBF and SMO then were measured using two performance criterions; accuracy and MSE. In general, by graphical observation, it can be seen that the plot of support vectors in hyperplane were different to each other and it is based on numbers of support vectors produced by RBF and SMO, accordingly (Refer to Figure 2 and Figure 3).

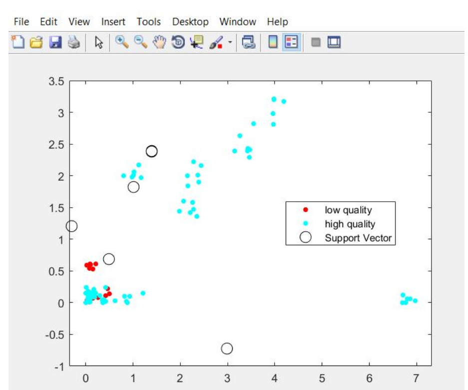

Figure 2: The plotted of support vectors using RBF training algorithm

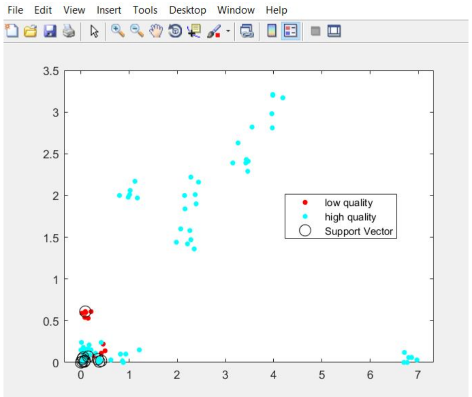

Figure 3: The plotted of support vectors using SMO training algorithm

As shown in the Table 3, the accuracy obtained by RBF and SMO algorithms are 94.7368 % and 89.4700 %, respectively. It was followed by the MSE for RBF which 0.0625 and for SMO which is 0.0000%. They showed that both accuracies are above 80 % and both MSE are close to 0.0000, indicate a good performance of modelling. It also proved that the analytical method by solving the two Lagrange multipliers in SMO and Gaussian function in RBF is able to grade the agarwood oil compounds successfully to their respective grade.

Table 3: The MSE or error rate for RBF and SMO testing

Algorithm	Accuracy (%)	Mean Square (MSE)
RBF	94.7368	0.0625
SMO	89.4700	0.0000

4. CONCLUSION

The study in this paper has successfully done and proven the successful of the agarwood oil quality grading using Radial Basis Function (RBF) and Sequential Minimal Optimization (SMO) in SVM Modelling. The analytical method by solving the two Lagrange multipliers in SMO and Gaussian function in RBF is able to grade the agarwood oil compounds successfully to their respective grade. This finding in this study is not just important, however, it also significant and benefits to the agarwood oil trading industry especially to its grading system and related research area.

ACKNOWLEDGEMENT

This research is funded by Ministry of Higher Education (MOHE) under FRGS Grant No.600-IRMI/FRGS 5/3 (031/2019) through Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia.The authors would like to thank Associate Professor Ts.Dr. Mohd. Hezri Fazalul Rahman for his valuable knowledge contribution in this article as well as all the ASPRG, FKE, FRIM and BARCE staff that involved in this study.

REFERENCES

[1] Aziz, M. A. A., Ismail, N., Yassin, I. M., Zabidi, A., & Ali, M. S. A. M. (2016). Agarwoodoil quality classification using cascade-forward neural network. In Proceedings - 2015 6th IEEE Control and System Graduate Research Colloquium, ICSGRC 2015 (pp. 112–115). Institute of Electrical and Electronics Engineers Inc.

[2] Liu, Y., Wei, J., Gao, Z., Zhang, Z., & Lyu, J. (2017). A Review of Quality Assessment and Grading for Agarwood. Chinese Herbal Medicines, 9(1), 22-30

[3] Ismail, N., Ali, N. A. M., Jamil, M., Rahiman, M. H. F., Tajuddin, S. N., & Taib, M. N. (2014). A review study of agarwood oil and its quality analysis. Jurnal Teknologi (Sciences and Engineering), 68(1), 37–42.

[4] Barden, A., Anak, N. A., Mulliken, T., & Song, M. (2000). Heart of the matter: agarwood use and trade and CITES implementation for Aquilaria malaccensis. TRAFFIC International, Cambridge, UK.

[5] Noble, W. S. (2006, December). What is a support vector machine? Nature Biotechnology.

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

[7] Trotter, M. W., Buxton, B. F., & Holden, S. B. (2001). Support vector machines in combinatorial chemistry. Measurement and Control, 34(8), 235-239.

[8] N. Cristianini and J. Shawe-Taylor, An introduction to support vector machines and other kernel-based learning methods: Cambridge university press, 2000.

[9] R. P. Lippman. "A Critical Overview of Neural Network Pattern Classifiers," Proceedings of the IEEE Neural Networks for Signal Processing Workshop, 1991, Princeton, N.J., pp. 266-275.

[10] R.P. Lippman, "Pattern Classification Using Neural Networks," IEEE Communications Magazine, 1989, pp. 47-64.

[11] S.J. Nowlan. Soft Competitive Adaptation: Neural Network Learning Algorithms based on Fitting Statistical Mixtures, Ph.D. thesis, CMU-CS-9 1- 126, Carnegie Mellon University, 1991.

[12] David P. Morgan and Christopher L. Scofield. Neural Networks and Speech Pro-processing, Kluwer Academic Publishers, Boston, 1991, p. 64.

[13] Donald F. Specht, "A General Regression Neural Network," IEEE Transactions on Neural Nenuorks, November 1991, Vol. 2, Number 6, pp. 568-576.

[14] Watkins, S. S., Chau, P. M., & Tawel, R. (1992, June). A radial basis function neurocomputer implemented with analog VLSI circuits. In [Proceedings 1992] IJCNN International Joint Conference on Neural Networks (Vol. 2, pp. 607-612). IEEE.

[15] Micchelli, C. A. (1986). Interpolation of scattered data: Distance matrices and conditionally positive definite functions. Constructive approximation. 2, 11-22.

[16] Z.-J. Fu, W. Chen, and L. Ling, "Method of approximate particular solutions for constant-and variable-order fractional diffusion models," Engineering Analysis with Boundary Elements, vol. 57, pp. 37-46, 2015.

[17] Park, J., & Sandberg, I. W. (1991). Universal approximation using radial-basis-function networks. Neural computation, 3(2), 246-257.
[18] Pavithra G, Abirami P, Bhuvaneshwari S, Dharani S and Haridharani B, A Survey on Intrusion Detectin Mechanism using Machine Learning Algorithms , Int. J. of Emerging Trends in Engineering Research (IJETER), Vol. 8, No. 4, 2020.
[19]. N. Chandra Sekhar Reddy, Purna Chandra Rao Vemuri and A. Govardhan, An Empirical Study on Support Vector Machine for Intrusion Detection, Int. J. of Emerging Trends in Engineering Research (IJETER), Vol. 7, No. 10, 2019.
[20] N. Ismail, "ANN Modelling of Agarwood Oil Significant Chemical Compounds for Quality Discrimination," 2013.