JOIV : Int. J. Inform. Visualization, 8(2) - May 2024 592-597

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Firefly Algorithm for SVM Multi-class Optimization on Soybean Land Suitability Analysis

Andi Nurkholis [a], Styawati [b], Alvi Suhartanto [c,*]

[a] *Department of Informatics, Universitas Teknokrat Indonesia, Kedaton, Bandar Lampung, Indonesia*
[b] *Department of Information Systems, Universitas Teknokrat Indonesia, Kedaton, Bandar Lampung, Indonesia*
[c] *Department of Information Technology, Universitas Teknokrat Indonesia, Kedaton, Bandar Lampung, Indonesia*
*Corresponding author: *alvi.suhartanto@teknokrat.ac.id*

***Abstract*— Soybean is the primary source of vegetable protein nutrition, containing fat and vitamins that Indonesian people widely consume. The decline in soybean production in Indonesia every year is due to the reduced area of soybean cultivation, thereby increasing dependence on imports from other countries. Land suitability maps can provide directions for priority locations for soybean cultivation based on land characteristics and weather to produce optimal production. The SVM multi-class algorithm has been applied to classify land suitability data to create a land suitability map but has yet to obtain optimal accuracy, especially for sigmoid kernels. The objective of this study is to enhance the performance of the sigmoid kernel SVM by utilizing the firefly algorithm. The study focuses on evaluating the suitability of soybean cultivation in Bogor and Grobogan Regencies. The results of the tests indicate that the firefly algorithm-optimized SVM (FA-SVM) significantly improves accuracy compared to the SVM without optimization. The accuracy achieved by FA-SVM is 89.95%, while the SVM without optimization only achieves an accuracy of 65.99%. The best parameters produced by the firefly algorithm are C=2.33 and σ=0.45 obtained from firefly customization, and the number of generations is 10. Based on this, the optimization algorithm can be used to produce an optimal model. The best optimal model obtained can be used as a guide for priority locations/areas for soybean cultivation by farming communities, so as to produce maximum soybean productivity.**

***Keywords*— Firefly; land suitability; sigmoid; soybean; SVM.**

Manuscript received 29 May. 2023; revised 2 Oct. 2023; accepted 20 Nov. 2023. Date of publication 31 May. 2024.

I. INTRODUCTION

Soybean, a legume plant, plays a vital role as a key ingredient in various processed foods including milk, soy sauce, tofu, and tempeh [1]. The shape of a soybean is similar to that of a pea, and a single plant can yield anywhere from 100 to 250 peas. In Indonesia, soybean serves as a primary source of vegetable protein, providing essential nutrition [2]. Moreover, soybean contains fats and vitamins that are beneficial for the body's well-being [3], [4]. Many processed food items and beverages extensively utilize soybeans as their primary component, such as tofu, tempeh, soy flour, soy milk, snacks, and soybean oil. Soybean oil is further processed to manufacture a range of products including soaps, plastics, inks, resins, solvents, crayons, and cosmetics [5].

The people of Indonesia heavily rely on soybeans for various purposes, highlighting the importance of ensuring an adequate supply of this crop. However, there is a concerning trend as the Ministry of Agriculture predicts a continued decline in soybean production in Indonesia until 2024 [6]. According to the projections for 2021, the anticipated domestic soybean production is 613.3 thousand tons, reflecting a decrease of 3.01% compared to the production of 632.3 thousand tons in 2020. Furthermore, there is a projection that forecasts a decrease of 3.05% in soybean production in Indonesia for the year 2022, with a total output estimated to be 594.6 thousand tons. The following year, a further decrease of 3.09% is predicted, resulting in a soybean production of 576.3 thousand tons. The Ministry of Agriculture attributes this decline to intense competition for land utilization with other strategic commodities like corn and chili, leading to a yearly reduction in harvested area of approximately 5%. To address this issue, one potential solution is the creation of a soybean land suitability map, which would offer guidance on optimizing land usage based on soybean planting conditions [7]–[9].

Land suitability maps can be generated through the assessment of the current suitability of land. This evaluation involves assessing the capacity of land resources based on the suitability of a specific agricultural commodity [10]. The aim of land suitability evaluation is to develop a predictive model that can be applied to other regions using land and weather characteristics to generate new land suitability maps [11]. The conventional approach for assessing land suitability typically involves comparing land and weather characteristics with the land suitability classes defined by the Food and Agriculture Organization (FAO) standards. These classes consist of S1 (very suitable), S2 (moderately suitable), S3 (marginally suitable), and N (not suitable) [12], [13].

This study builds upon previous research that focused on utilizing the Support Vector Machine (SVM) Multi-Class classification approach to assess the suitability of soybean cultivation in Bogor and Grobogan Regencies [14]. That research resulted in six different model variations based on the combination of kernels and the number of fold cross-validation (CV). The highest accuracy reached 96.91%, built using the RBF kernel, while the lowest accuracy was 65.99%, made using the sigmoid kernel. The significant difference between the results of the two kernels is a problem that needs to be resolved. One solution that can be proposed is the implementation of an optimization algorithm on the sigmoid kernel to find out whether the optimization algorithm has an impact and can improve its accuracy. The algorithm that can be used for SVM optimization is Firefly which, based on previous research, has a significant effect compared to other optimization algorithms [15], [16].

This study aims to optimize the results of the SVM Multi-Class model based on the sigmoid kernel in evaluating soybean land suitability using the firefly algorithm. The firefly algorithm parameters will also be analyzed to obtain the best accuracy. The optimization of the model is anticipated to yield highly accurate outcomes, ultimately contributing to the creation of an improved soybean land suitability map.

II. Material and Method

The research study focuses on two regencies, namely Bogor (West Java Province) and Grobogan (Central Java Province), with respective land areas of 299,070 hectares (ha) [17] and 202,867 ha [18]. Bogor Regency was selected based on a previous study on soybean land adaptability, which achieved a sufficiently accurate model [7], indicating its potential as a reference for optimal soybean land adaptation. With a percentage of 43.08% [19], Grobogan Regency plays a crucial role in soybean production within Central Java Province and is expected to serve as a benchmark for establishing the best land suitability standards in other regions. The data from these two regencies were combined to create a more extensive and comprehensive dataset. The study utilized two types of data: the explanatory and the target class. The explanatory consisted of nine factors, including two meteorological variables from BMKG and seven land characteristics from BBSDLP. The target class, representing soybean land suitability, was based on the BBSDLP mapping from previous studies. Each attribute used in the study is described in Table 1.

TABLE I
RESEARCH DATA

Attribute	Description	Format	Source
Drainage*	The categorization of how water impacts the level of soil aeration in different areas.	Vector	BBSDLP
Land slope (%)	The incline percentage of the terrain.	Vector	BBSDLP
Soil pH (°)	The nutritional content of the soil.	Vector	BBSDLP
Soil texture*	The categorization of terms regarding the dispersion of small soil particles measuring 2 mm in size.	Vector	BBSDLP
Cation exchange capacity (cmol)	The numerical value representing the ability of the clay fraction to exchange cations.	Vector	BBSDLP
Base saturation (%)	The number of bases (NH_4OAc) present in 100g of soil sample	Vector	BBSDLP
Depth of soil mineral (cm)	The depth measurement of minerals within the soil layer.	Vector	BBSDLP
Rainfall (mm)	The cumulative amount of rainfall during a month (October 2019)	Spreadsheet	BMKG
Temperature (°C)	The mean temperature value during a month (October 2019)	Spreadsheet	BMKG
Soybean land suitability	The soybean land suitability class: S1, S2, S3, and N	Vector	BBSDLP

*Attributes have no quantitative value

This work involved three main phases: data preprocessing, optimization of the SVM model using the Firefly algorithm, and visualization of the soybean land suitability map. The process flow of these steps is illustrated in Figure 1.

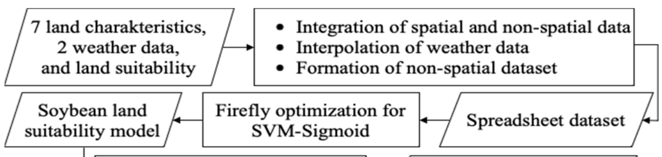

Fig. 1 Steps of the study

The following provides a thorough overview of the research steps based on Fig. 1.

A. Data Preprocessing

This research is a development of previous research, where the dataset used is the same. In previous research, three stages of data preprocessing have been carried out, which aim to produce the final dataset, namely non-spatial data in spreadsheet format. As it is known, the initial dataset contains several spatial attributes in vector format, especially land characteristics. Based on this, it is necessary to preprocess the data to produce datasets in a format suitable [20] for modeling using FA-SVM, namely a non-spatial dataset.

In the initial step of data preprocessing, the integration of spatial and non-spatial data from the Grobogan Regency BBSDLP was performed. This dataset included various variables such as drainage, relief, base saturation, cation exchange capacity, soil texture, soil pH, depth, and soil minerals. The data was available in two formats: spatial objects in vector format and non-spatial attributes in spreadsheet format. To combine these datasets, they were merged based on the soil map unit (SPT), which served as a unique identifier linking a data line to a spatial object. In this study, each SPT could represent one or multiple spatial objects in polygon form. The merging process was executed using PostgreSQL Database Management System (DBMS) version 13.1. Consequently, seven layers of land characteristics were created, each attached to its respective spatial object as non-spatial properties.

The second stage of data preprocessing involved interpolating weather data to generate rainfall and temperature layers. Interpolation was necessary, and Ordinary Cokriging (OCK) was chosen for its higher accuracy compared to other techniques such as Ordinary Kriging (OK) and Kriging with External Drift (KED) [21]. For rainfall data interpolation, the primary variable utilized was the total precipitation amount for a specific month, while the elevation value served as the supporting variable. In the case of temperature data interpolation, the primary variable was the average monthly temperature, with the elevation value as the supporting variable. The resulting temperature and rainfall layers were in raster format, comprising pixel points representing the distribution values based on the closest meteorological station. The "Add Surface Information" tool in the ArcMap software was then employed to assign these points to their respective polygons based on the average values. This insertion process resulted in each polygon having ten non-spatial attributes, including a target class, nine explanatory factors, and one additional attribute.

The last stage of data preprocessing involved the transformation of spatial data represented by polygons, which were of vector type. In this study, each data row was divided based on polygons, resulting in non-spatial variables that included nine explanatory factors and a target class. For instance, if a regency had ten polygons, it would produce a non-spatial dataset consisting of ten rows of data. The utilization of polygons to separate the data was crucial because it allowed for the consideration of potential variations in rainfall and temperature values within each polygon, even when the same soil map unit (SPT) was used.

B. *Firefly Algorithm for SVM*

SVM (Support Vector Machine) is widely acknowledged as a highly effective supervised machine learning algorithm [22]. Initially, SVM was primarily designed to handle binary classification tasks [23]. However, in this study, there were four land suitability classes, namely S1, S2, S3, and N. To accommodate multi-class scenarios where k>2, SVM can be modified using either the One Against One or One Against Rest method [24].

Both the One Against One and One Against Rest methods have their own strengths and weaknesses, but the One Against One approach is considered superior due to its computational efficiency and lower complexity [25]. Previous studies have also compared these two methods in sentiment analysis tasks, and they have reported similar or identical accuracies [26]. In this study, the One Against One method was specifically employed using the e1071 library in the R programming version 4.0.3. This approach entails identifying $k(k-1)/2$ separator functions, where each function is trained using data from two classes. The training data is denoted as x_i, y_i, where i varies from 1 to n, with n representing the total number of data points. The available data is denoted as $\vec{x_i} \in R^d$, where $\{x_i = x_{i1}, x_{i2}, x_{i3}, \dots, x_{iq}\}^T$ represents the attributes or features of the i-th training data. The labels or classes are represented as $y_i \in \{-1, +1\}$, where i spans from 1 to n. It is assumed that these two classes, -1 and $+1$, can be effectively separated by a d-dimensional hyperplane using the kernel trick concept in a higher-dimensional space. This research builds upon previous studies that obtained low accuracy when using sigmoid kernel parameters in SVM classification for soybean land suitability analysis.

In previous research [14], the sigmoid kernel was applied to SVM without adjusting parameters, which means using standard settings. This allows the parameter settings used to be less than optimal in the sigmoid kernel, which causes the accuracy obtained to be low. This is supported by research, which states that determining the parameters C and σ has an influence, where if the values of these two parameters are by the kernel characteristics, optimal results will be obtained and the opposite [15]. To address this issue, the firefly algorithm is implemented for SVM optimization, commonly referred to as FA-SVM, to improve the accuracy of the results.

The FA-SVM classification process commences by initializing crucial parameters required for the search process through the use of the firefly algorithm [16]. These parameters include the firefly numbers, the generation numbers, the initial attractiveness coefficient (β_0), the light absorption coefficient (γ), and the random parameter coefficient (α). Once the parameters have been initialized, the firefly algorithm is employed to optimize the values of C and σ. The optimization process involves the following steps [27]:

1. Set objective function $f(x)$
2. Set firefly population $(i = 1,2, \dots N)$
3. Set light absorption coefficient γ.
4. Count distance between two fireflies, i and j, located at coordinates x_i and x_j in Cartesian space using the following formula:

$$r_{ij} = \|m_i - m_j\| = \sqrt{\textstyle\sum_{k=1}^{d} \left(m_{i,k} - m_{j,k}\right)^2} \tag{1}$$

If the equation involves two dimensions $(d = 2)$, the aforementioned equation can be expressed as [28]:

$$r_{ij} = \sqrt{\left(m_i - m_j\right)^2 + \left(n_i - n_j\right)^2} \tag{2}$$

1. Calculate the attraction of a firefly, which is inversely related to the brightness of the light it can perceive. The

attractiveness can be calculated using the following formula:

$$\beta(r) = \beta_0 e^{-\gamma r^2} \quad (3)$$

where, $\beta(r)$ represents the attractiveness of fireflies at a distance r, β_0 denotes the attractiveness at a distance of 0, γ is the coefficient of light absorption, and r indicates the distance between the source fireflies and the target fireflies.

2. Calculate the movement of fireflies that are attracted to fireflies j, which are brighter or more appealing, using the following formula:

$$m_i = m_i + \beta_0 e^{-\gamma r_{ij}^2}\left(m_j - m_i\right) + \alpha\left(rand - \frac{1}{2}\right) \quad (4)$$

Basically, the parameter values employed are and $\alpha \in [0,1]$. The randomization procedure can be implemented by utilizing a suitable distribution, such as a normal distribution like $N(0,1)$, or any other appropriate distribution.

Once the values of C and σ are obtained from each firefly, they are employed to train the data using the SVM method. Subsequently, the accuracy of each C and σ produced by the fireflies is computed. The accuracy results are then ranked to identify the most optimal values for the parameters C and σ. These optimized values are utilized to construct the SVM classifier model. Finally, the accuracy of the model is evaluated by testing it on a separate dataset to assess its performance in predicting the outcomes of the test data. This study will optimize the sigmoid kernel using the FA-SVM classification. This is based on the low accuracy produced in previous research, so it is necessary to do optimization to increase accuracy and create an optimal model. The sigmoid kernel function formulated as follows:

$$K\left(x_i, x_j\right) = \tanh\left(\sigma\left(x_i, x_j\right) + c\right) \quad (5)$$

To evaluate and enhance the performance of the model, the accuracy metric is employed to assess the model's effectiveness in predicting the data accurately when compared to the actual data. Higher accuracy values indicate lower prediction errors in the test data, indicating improved model performance. In this study, the approach of k-fold cross-validation is employed. This approach entails randomly partitioning the sample set into k subsets, which is repeated k times. In each iteration, one subset is used for testing, while the remaining subsets are utilized for training [29], [30]. The accuracy of the classification model is determined by comparing the test data with the model's predictions, using Equation 6 as specified in reference [31].

$$Accuracy(\%) = \frac{\sum Test\ data\ is\ correctly\ classified}{\sum Test\ data} \times 100 \quad (6)$$

C. *Visualization*

The soybean land suitability map for the Bogor and Grobogan regencies will be generated based on the prediction results of the best-performing model in this research study. The ArcMap application is employed to create a spatial map, which plays a crucial role in the visualization [32].

III. Result and Discussion

After the completion of the data preprocessing stage, a dataset comprising a total of 388 data points was obtained. The dataset utilized in this research comprises 238 rows from Bogor Regency and 150 rows from Grobogan Regency. The combined dataset encompasses ten characteristics, consisting of nine explanatory factors and one target class.

A. *FA-SVM classifier for Soybean Land Suitability*

The FA-SVM modeling process commences by initializing the necessary parameters for the process using firefly optimization. These parameters include the population size of fireflies, the number of generations, the initial attractiveness, the light absorption, and the random parameter. The next step involves optimizing the C and σ parameters obtained from each firefly. The accuracy of each C and σ value generated by the fireflies is then calculated and ranked to determine the optimal parameter values. These optimized values are utilized to construct the SVM classifier model. In order to evaluate the performance of the model, a K-Fold cross-validation technique is utilized. Ten folds are used in this study, as previous research has shown that it leads to higher accuracy compared to using five folds. The FA-SVM modeling is implemented using the R programming version 4.0.3 and the e1071 library. Furthermore, the models are created by adjusting the firefly and generation numbers in order to optimize the accuracy achieved.

B. *Model Evaluation*

The performance evaluation of the FA-SVM method with a sigmoid kernel was conducted using 10-fold cross-validation, ensuring equal-sized and randomly partitioned data subsets. Each fold involved dividing the data into ten subsets, with each subset having the same size but containing different data instances. For instance, in a dataset of 388 data points, one-fold would consist of 39 instances as the testing data and 349 instances as the training data. This process was repeated ten folds, resulting in different distributions of data. Prior to evaluating the model, the Firefly method was employed to search for the optimal SVM parameters. The sought SVM parameters fall within the range of C=1.0-3.0, σ=0.1-1.0, as indicated by previous research as the optimal range for SVM parameters. Table 2 provides a comparison of accuracy based on different customization options for the number of fireflies, while Table 3 presents a comparison of accuracy based on various customization options for the number of generations.

TABLE II
SVM-FA MODEL EVALUATION BASED ON NUMBER OF FIREFLY

Number of Firefly	Number of Generation	C	σ	Accuracy (%)
10	10	2.33	0.45	89.95
20	10	2.25	0.46	89.69
30	10	2.13	0.50	89.95

TABLE III
SVM-FA MODEL EVALUATION BASED ON NUMBER OF GENERATION

Number of Firefly	Number of Generation	C	σ	Accuracy (%)
10	10	2.33	0.45	89.95
10	20	2.28	0.46	89.95
10	30	2.20	0.47	89.95

Table 2 presents the results showing that the models customized with 10 and 30 fireflies achieved the highest levels of accuracy. Interestingly, the model with 20 fireflies also yields comparable accuracy, differing by only 0.26%. The difference in the number of fireflies does not affect the resulting accuracy. This is also in line with the results of the comparison of accuracy based on the number of generations in Table 3, where the difference in the number of generations of the model still produces the same accuracy, namely 89.95%. Unlike the case with accuracy, the C and σ values generated in each model are different but not significant. Based on the test results, it can be said that the customization of the number of fireflies and generation will only slightly affect the value of C and σ but produces almost the same accuracy. Following previous research [14], applying the firefly algorithm for sigmoid kernel optimization in the SVM classification obtained positive results, which increased accuracy from initially only 65.99% to 89.95%. It can be concluded that setting the parameters C and σ in the SVM classification greatly affects the accuracy results obtained, where the firefly algorithm can provide the correct values to produce optimal accuracy.

This increase in accuracy proves that sigmoid kernel optimization in SVM can be carried out effectively by the firefly algorithm. The firefly algorithm influences specific parameter settings to produce the most optimal combination. This means that the sigmoid kernel in SVM will work better when it gets the proper parameter settings or combinations. Based on this explanation, this research successfully corrected the weaknesses in the previous study [14] to produce much better accuracy through kernel parameter optimization by the firefly algorithm.

C. Soybean Land Suitability Maps

Moreover, the optimal model outcomes were visually represented as land suitability maps for Bogor Regency and Grobogan Regency. Fig. 2 illustrates the predicted land suitability mapping results for soybeans, allowing for a comparison with the actual data version provided by BBSDLP.

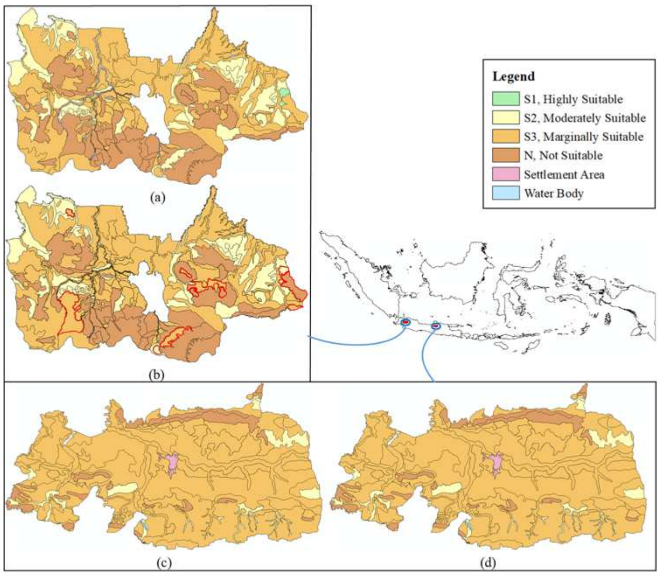

Fig. 2 Maps of Soybean Land Suitability of (a) BBSDLP and (b) Sigmoid in Bogor, (c) BBSDLP and (d) Sigmoid in Grobogan

Fig. 2 demonstrates a noticeable inconsistency in the soybean land suitability maps. Specifically, the model version of the map for Bogor Regency (Fig. 2(b)) lacks the S1 class, which is present in the BBSDLP version (Fig. 2(a)). This discrepancy is evident from the red polygon boundary, which is derived from a different model than the actual data version. The lack of the S1 class in the model's results can be attributed to the limitations of the model. However, in the case of Grobogan Regency, both the model version (Fig. 2(c)) and the BBSDLP version (Fig. 2(d)) of the land suitability map demonstrate coherence and consistency. This indicates that the model can effectively predict soybean land suitability in Grobogan Regency. The PostgreSQL, particularly the ST Area function, was employed to compute the area of each land suitability class, offering valuable information on the suitability of land for soybean cultivation. Table 3 presents the directions for soybean land use in both Bogor and Grobogan regencies, based on the information obtained from both the BBSDLP version and the model.

TABLE IV
SOYBEAN LAND SUITABILITY AREA

Land suitability class	**Area Total (ha)**			
	Bogor		**Grobogan**	
	BBSDLP	**Sigmoid Model**	**BBSDLP**	**Sigmoid Model**
S1	881.48	-	-	-
S2	53,069.2	52,290	10,697.27	10,697.27
S3	150,165.2	158,863.5	180,365.14	180,365.14
N	93,963.47	86,925.81	15,227.66	15,227.66
Settlement area	-	-	1,033.64	1,033.64
Water body	943.46	943.46	104.09	104.09

Table 3 presents a comparison between the results of the model and the BBSDLP version regarding the classification of soybean land suitability. The main focus is on the discrepancies observed in Bogor Regency. The model's findings reveal several differences, such as the absence of the S1 class and modifications in the classification of soybean land suitability in other classes. This indicates that the model's accuracy of 89.95% suggests that the model may not accurately identify all data points. Consequently, the dominant soybean land suitability classes in both Bogor and Grobogan Regencies are identified as S3, N, and S2. This underscores the importance of careful planning for soybean production, with particular attention given to areas classified as S1, S2, and S3 in Bogor and Grobogan Regencies.

IV. CONCLUSION

This study succeeded in optimizing the accuracy of the sigmoid kernel in the SVM classification using the firefly algorithm, which reached 89.95%. FA-SVM significantly improves the performance of the soybean land suitability prediction model's performance in the Bogor and Grobogan Regencies case studies. The best parameters produced by firefly optimization are C=2.33 and σ=0.45 obtained through a combination of 10 fireflies and generation. The evaluation results show that the difference in the number of fireflies and generation (minimum 10) does not significantly affect the accuracy obtained but can produce a value C and σ are different. The resulting land suitability map can guide related parties (farmers, the private sector, and local government) to prioritize soybean cultivation in S1, S2, and S3 class areas to increase soybean productivity. As a development, further research can further analyze other optimization algorithms to produce better accuracy. Furthermore, as part of the implementation of the model results, it is possible to develop a geographic information system (GIS) specifically designed for soybean land suitability. This GIS can be made available online, providing a mapping platform that can be accessed by the community, thus increasing its usefulness and accessibility.

ACKNOWLEDGMENT

This research was funded by *Hibah Publikasi Terindeks Scopus* (HIPUTS) Universitas Teknokrat Indonesia, Contract Number: 020/UTI/LPPMI/E.1.1/VIII/2023

REFERENCES

[1] R. Krisdiana, N. Prasetiaswati, I. Sutrisno, F. Rozi, A. Harsono, and M. J. Mejaya, "Financial Feasibility and Competitiveness Levels of Soybean Varieties in Rice-Based Cropping System of Indonesia," *Sustainability*, vol. 13, no. 15, p. 8334, 2021.

[2] A. Romulo and R. Surya, "Tempe: A traditional fermented food of Indonesia and its health benefits," *Int J Gastron Food Sci*, vol. 26, p. 100413, 2021.

[3] L. Liu *et al.*, "Traditional fermented soybean products: processing, flavor formation, nutritional and biological activities," *Crit Rev Food Sci Nutr*, vol. 62, no. 7, pp. 1971–1989, 2022.

[4] P. Qin, T. Wang, and Y. Luo, "A review on plant-based proteins from soybean: Health benefits and soy product development," *J Agric Food Res*, vol. 7, p. 100265, 2022.

[5] A. C. Akmalovna, "Characteristics and Advantages of Soybean Benefits in Every way," *Journal of Ethics and Diversity in International Communication*, vol. 1, no. 8, pp. 67–69, 2022.

[6] Pusat Data dan Sistem Informasi Pertanian, *Outlook Komoditas Pertanian Tanaman Pangan Kedelai*. Kementerian Pertanian, 2020.

[7] A. Nurkholis and I. S. Sitanggang, "A spatial analysis of soybean land suitability using spatial decision tree algorithm," Sixth International Symposium on LAPAN-IPB Satellite, Dec. 2019, doi:10.1117/12.2541555.

[8] A. Nurkholis and S. Styawati, "Prediction Model for Soybean Land Suitability Using C5. 0 Algorithm," *Jurnal Online Informatika*, vol. 6, no. 2, pp. 163–171, 2021.

[9] P. Munene, L. M. Chabala, and A. M. Mweetwa, "Land Suitability Assessment for Soybean (Glycine max (L.) Merr.) Production in Kabwe District, Central Zambia," Journal of Agricultural Science, vol. 9, no. 3, p. 74, Feb. 2017, doi: 10.5539/jas.v9n3p74.

[10] A. Nurkholis, I. S. Sitanggang, A. Annisa, and S. Sobir, "Spatial decision tree model for garlic land suitability evaluation," IAES International Journal of Artificial Intelligence (IJ-AI), vol. 10, no. 3, p. 666, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp666-675.

[11] I. S. Sitanggang, A. Nurkholis, . Annisa, and M. A. Agmalaro, "Garlic Land Suitability System based on Spatial Decision Tree," Proceedings of the International Conferences on Information System and Technology, 2019, doi: 10.5220/0009908002060210.

[12] A. K. Nisyak, F. Ramdani, and Suprapto, "Web-GIS development and analysis of land suitability for rice plant using GIS-MCDA method in Batu city," 2017 International Symposium on Geoinformatics (ISyG), Nov. 2017, doi: 10.1109/isyg.2017.8280667.

[13] I. W. Nuarsa, I. N. Dibia, K. Wikantika, D. Suwardhi, and I. N. Rai, "Gis based analysis of agroclimate land suitability for Banana plants in Bali Province, Indonesia," *Hayati*, vol. 25, no. 1, pp. 11–17, 2018.

[14] A. Nurkholis, S. Styawati, I. S. Sitanggang, J. Jupriyadi, A. Matin, and P. Maulana, "SVM Multi-Class Algorithm for Soybean Land Suitability Evaluation," in *2022 International Conference on Information Technology Research and Innovation (ICITRI)*, IEEE, 2022, pp. 65–70.

[15] S. Styawati and K. Mustofa, "A Support Vector Machine-Firefly Algorithm for Movie Opinion Data Classification," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 13, no. 3, pp. 219–230, 2019.

[16] H. Zhang, Y. Shi, X. Yang, and R. Zhou, "A firefly algorithm modified support vector machine for the credit risk assessment of supply chain finance," *Res Int Bus Finance*, vol. 58, p. 101482, 2021.

[17] BBSDLP, *Atlas peta kesesuaian lahan dan arahan komoditas pertanian pertanian, Kabupaten Grobogan, Provinsi Jawa Tengah, skala 1:50.000*. Bogor (ID): Badan Penelitian dan Pengembangan Pertanian, Kementerian Pertanian, 2016.

[18] BBSDLP, *Atlas peta kesesuaian lahan dan arahan komoditas pertanian pertanian, Kabupaten Bogor, Provinsi Jawa Barat, skala 1:50.000*, 2nd ed. Bogor (ID): Badan Penelitian dan Pengembangan Pertanian, Kementerian Pertanian, 2016.

[19] Pusdatin, *Outlook Kedelai: Komoditas Pertanian Subsektor Tanaman Pangan*. Jakarta (ID): Pusat Data dan Sistem Informasi Pertanian Kementerian Pertanian, Kementerian Pertanian RI, 2018.

[20] A. Nurkholis, Styawati, D. Alita, A. Sucipto, M. Chanafy, and Z. Amalia, "Hotspot Classification for Forest Fire Prediction using C5.0 Algorithm," in *International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*, Bandung: IEEE, 2021.

[21] S. K. Adhikary, N. Muttil, and A. G. Yilmaz, "Cokriging for enhanced spatial interpolation of rainfall in two Australian catchments," Hydrological Processes, vol. 31, no. 12, pp. 2143–2161, Apr. 2017, doi: 10.1002/hyp.11163.

[22] E. J. Camp *et al.*, "Supervised machine learning to predict reduced depression severity in people with epilepsy through epilepsy self-management intervention," *Epilepsy & Behavior*, vol. 127, p. 108548, 2022.

[23] P. Sahoo, A. K. Behera, M. K. Pandia, C. S. K. Dash, and S. Dehuri, "On the study of GRBF and polynomial kernel based support vector machine in web logs," in *2013 1st International Conference on Emerging Trends and Applications in Computer Science*, IEEE, 2013, pp. 1–5.

[24] D. Alita, "Multiclass Svm Algorithm For Sarcasm Text In Twitter," *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, vol. 8, no. 1, pp. 118–128, 2021.

[25] D. Alita, S. Priyanta, and N. Rokhman, "Analysis of emoticon and sarcasm effect on sentiment analysis of Indonesian language on Twitter," *Journal of Information Systems Engineering and Business Intelligence*, vol. 5, no. 2, pp. 100–109, 2019.

[26] A. Nurkholis, D. Alita, and A. Munandar, "Comparison of Kernel Support Vector Machine Multi-Class in PPKM Sentiment Analysis on Twitter," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 2, Apr. 2022.

[27] X.-S. Yang, Z. Cui, R. Xiao, A. H. Gandomi, and M. Karamanoglu, *Swarm intelligence and bio-inspired computation: theory and applications*. Newnes, 2013.

[28] B. Xing and W.-J. Gao, *Innovative computational intelligence: a rough guide to 134 clever algorithms*, vol. 62. Springer, 2014.

[29] T. H. Kerbaa, A. Mezache, and H. Oudira, "Model Selection of Sea Clutter Using Cross Validation Method," Procedia Computer Science, vol. 158, pp. 394–400, 2019, doi: 10.1016/j.procs.2019.09.067.

[30] H. Ling, C. Qian, W. Kang, C. Liang, and H. Chen, "Combination of Support Vector Machine and K-Fold cross validation to predict compressive strength of concrete in marine environment," *Constr Build Mater*, vol. 206, pp. 355–363, 2019.

[31] A. Nurkholis, Muhaqiqin, and T. Susanto, "Algoritme Spatial Decision Tree untuk Evaluasi Kesesuaian Lahan Padi Sawah Irigasi," Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 4, no. 5, pp. 978–987, Oct. 2020, doi: 10.29207/resti.v4i5.2476.

[32] A. Nurkholis, M. Muhaqiqin, and T. Susanto, "Analisis Kesesuaian Lahan Padi Gogo Berbasis Sifat Tanah dan Cuaca Menggunakan ID3 Spasial," JUITA: Jurnal Informatika, vol. 8, no. 2, p. 235, Nov. 2020, doi: 10.30595/juita.v8i2.8311.