International Journal of Advances in Data and Information Systems
Vol. 5, No. 1, April 2024, pp. 1~11
ISSN: 2721-3056, DOI: 10.59395/ijadis.v5i1.1317

Data-Driven Analytical Model Using Machine Learning Algorithms: A Case Study on Clean and Healthy Living Behaviour in Surabaya City's Coastal Areas

Harun Al Azies[1,2*], Noval Ariyanto[1], Ishak Bintang Dikaputra[1]
[1]Study Program in Informatics Engineering, Faculty of Computer Science, Universitas Dian Nuswantoro, Indonesia
[2]Research Center for Materials Informatics, Faculty of Computer Science, Universitas Dian Nuswantoro, Indonesia

Article Info

Article history:

Received Sep 09, 2023
Revised Oct 31, 2023
Accepted Mar 18, 2024

Keywords:

Clean and Healthy Living Behavior
Machine learning
Support Vector Machine
Supervised Learning
Public Health

ABSTRACT

The objective of this article is to use machine learning technology, specifically the Support Vector Machine (SVM) approach with a linear kernel, to analyze and predict clean and healthy living behavior (CHLB) in coastal dwellings in Surabaya City. To train the SVM model, researchers collect health and environmental data from the region. As a result, our model predicts house CHLB status with an 83% accuracy rate. The most important variables in this prediction are the amount of community access to appropriate sanitary facilities, the health of households, and the sustainability of public areas that meet health requirements. These findings have crucial implications for attempts to improve CHLB in Surabaya's coastal areas in compliance with the National Medium-Term Development Plan (RPJMN) aims. Furthermore, the findings of this study can be used to build more targeted and long-term health policies in coastal communities.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Harun Al Azies,
Study Program in Informatics Engineering,
Universitas Dian Nuswantoro,
127 Imam Bonjol Street, Semarang 50131, Indonesia.
Email: harun.alazies@dsn.dinus.ac.id

1. INTRODUCTION

Surabaya's coastal region, which is part of East Java Province, has distinct characteristics for public health issues that necessitate specific care [1]. The region's location at the crossroads of immense economic potential, particularly in the trade and maritime sectors, and unique and complicated public health concerns provide a landscape that necessitates a creative and comprehensive strategy [2]–[4]. One of the most pressing issues that have to be addressed is the implementation of clean and healthy living behaviors among coastal households [5], [6]. CHLB is a crucial foundation for increasing the community's quality of life and welfare [7]. However, when compared to other locations, its implementation in Surabaya City's coastal zones has different dynamics and obstacles. CHLB stands out in Surabaya City's shoreline area due to the complexity of the factors that play a role [8], [9]. Coastal areas have significant cultural differences that influence health attitudes and practices. Inadequate access to health resources, particularly in coastal rural populations, is another issue that must be addressed. Mangroves, rocky beaches, and wetlands, for example, all play diverse roles in a household's CHLB status. Given these challenges, suitable and efficient approaches for completely determining the primary components that generate CHLB are required.

Previous studies on CHLB typically focused on conventional techniques to determine CHLB status and affecting variables. For instance, Yuni Susilowati's study on the Kurassaki school trash reduction campaign at BAPPEDA Tangerang Regency discovered that the program had a favorable impact on students' CHLB at BAPPEDA Tangerang Regency pilot project schools[10]. Hasyim's community empowerment research at Madrasah Diniyyah South Sumatra, conducted through CHLB, was successful in improving students' knowledge, attitudes, and behaviors. Through KIE, knowledge evaluations, and hand-washing demonstrations, this research also provides benefits such as creating a clean school environment, improving the teaching and learning process, and making students, instructors, and the school environment healthier as a whole. Finally, the need to promote clean and healthy living in schools is emphasized [11]. In addition, research by Ridha Rachmathiany at Al-Huda Islamic Boarding School in Kediri City that identified factors that affect Santri's CHLB reveals that knowledge of CHLB among Santri is still minimal and calls for increased attention from Islamic boarding school managers regarding CHLB [12]. Comparably Khairiyati et al. conducted an observational study in Sungai Paring Village on the factors that influence CHLB implementation at the household level. The findings of their study revealed that there was no significant relationship between knowledge, attitudes, and distance to healthcare facilities and CHLB implementation at the household level, but support from health workers and community leaders was related to CHLB implementation [13]. These findings illustrate the variety and complexity of factors impacting CHLB in different locales. This research has provided valuable insights into public health issues, but they have limits in extracting deeper information from available data and identifying the complex components that drive CHLB. To address this constraint, the new technique proposed by the researchers in this study is the application of machine learning technology. Researchers will be able to identify complex patterns and essential factors that traditional methods of analyzing and predicting CHLB status can overlook. As a result, this study will provide a more comprehensive and efficient approach to estimating the CHLB status of families in Surabaya City's coastal districts.

The machine learning approach proposed by the researchers in this work adds an entirely new dimension to comprehending and predicting CHLB status. This means that researchers will use artificial intelligence to study environmental data, community behavior, and other aspects never before considered in the context of CHLB in Surabaya City's coastline area. As a result, this study will provide a more comprehensive and efficient approach to predicting household CHLB status. This study offers the potential to establish better-focused and sustainable environmental health initiatives and health programs in Surabaya City's coastal districts. Our suggested unique strategy based on machine learning has the potential to significantly improve our understanding and treatment of CHLB in coastal locations, and it is a much-needed innovative step in addressing increasingly complicated public health concerns. Furthermore, in Part II, this article is a research technique used to describe the research stages, and it is utilized in Part III, namely Results and Discussion, to support the analysis of the results. Part IV is the research conclusion, which contains crucial information derived from a series of research findings.

2. RESEARCH METHOD

Figure 1 explains the research's conceptual framework. Data on CHLB were gathered from homes in Surabaya's coastline region as part of the process's initial data collection phase. This information serves as the basis for the entire study. This study uses secondary data, the data obtained from the Publication of the Surabaya City Health Office. Data about the public health profile of the city of Surabaya in the form of data related to the behavior of people's lives and the environmental conditions of the people of the city of Surabaya. The data used is Surabaya City Health profile data in the 2016-2018 period. The object of observation in this study is the community health center (Puskesmas) which represents households in the sub-district in the city of Surabaya which is included in the coastal area. A total of 24 health centers from 12 sub-districts which are coastal areas, the area that is the subject of this research is visualized in Figure 2, and the details of the health centers in each sub-district are presented in Table 1.

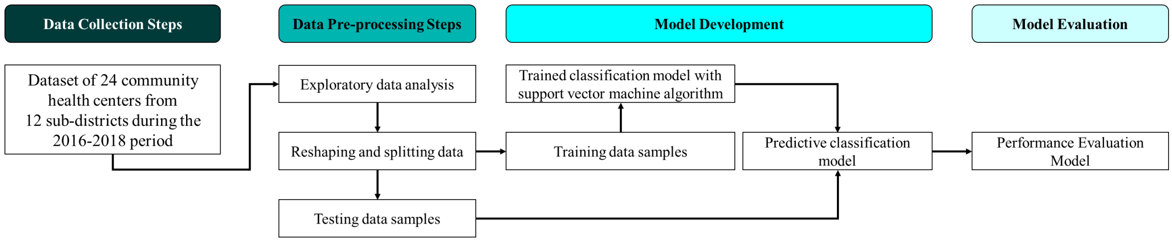


Figure 1. Conceptual framework for research on CHLB status classification of coastal households in Surabaya

Research variables consist of response variables and predictor variables. The household status with CHLB is the response or target variable. On a nominal scale, households can be classified as having good CHLB (percentage of households with CHLB 80%) or having less CHLB (percentage of households with CHLB 80%). Household CHLB assessment of achievement of performance indicator targets for community empowerment and health promotion activities based on RPJMN 2015 – 2019 (Rohman, 2020). The objective of the RPJMN for the percentage of areas with CHLB policies is 80%. The predictor variables used were five variables. The selection of several predictor variables is based on data on the state of the community environment in the coastal city of Surabaya. The variables include the percentage of residents with healthy homes (X_1), Percentage of the population with access to quality drinking water that meets health requirements (X_2), Percentage of the population with access to adequate sanitation facilities (X_3), Percentage of public facilities that meet health requirements (X_4), Percentage of food management establishments that meet sanitation hygiene status (X_5). Following that, we ran a data inquiry throughout the data preparation process. We used descriptive statistics to analyze the data [14]to acquire a general picture of house circumstances along the Surabaya coastline. The goal of this step is to have a preliminary, in-depth understanding of the data we have. The data is separated into two subsets after data exploration: training data and test data [15], [16]. We employ a commonly utilized technique in which training data constitutes 75% of total data and test data constitutes 25% of total data[17], [18]. This is an important stage since the training data will be used to train our model, while the test data will be used to assess the model's performance.

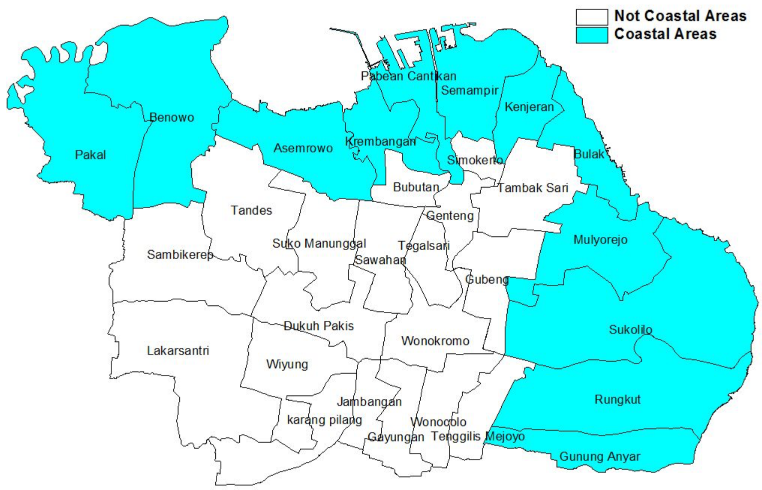
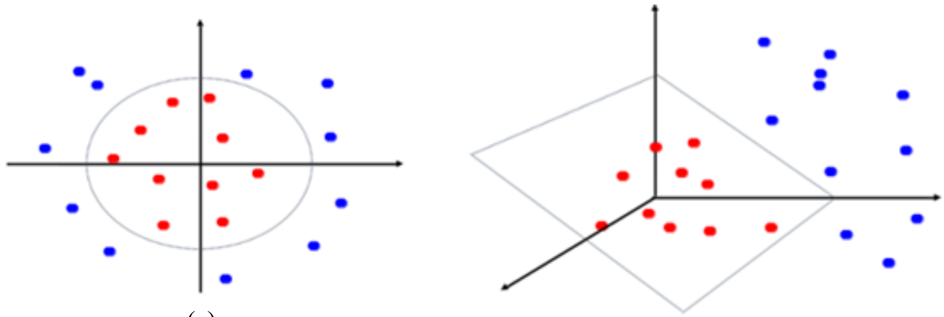


Figure 2. Research Location for Coastal Area in Surabaya City

The SVM method uses a supervised learning technique whose technical work is to classify[19]–[21]. Comparison with other classification methods, the SVM method is more powerful mathematically in its classification modeling [22], [23]. In addition, SVM can also solve classification and regression problems with linear and non-linear [24]. SVM is used to find the best hyperplane by maximizing the distance between classes[25]. A hyperplane is a function that can be used to separate classes [26]. In two dimensions the function used for classification between classes is called line, while the function used for classification between classes in three dimensions is called plane in the same way, while the function used for classification in a higher dimensional class space is called a hyperplane, make it easier to understand [27], as an illustration of how SVM works, it is visualized as follows.



(a) (b)

Figure 3. Illustration of Hyperplane, (a) 2-dimensional function, (b) 3-dimensional function [28]

The SVM method has three approach models or commonly called kernels, namely radial, linear and sigmoid kernels [29], [30]. In this study, the application of SVM using a linear kernel function will technically be explained in the analysis phase. Following the previous step, when we obtained the classification results, we moved on to the data classification stage. To assess the SVM method's accuracy, we examined the data using a model that we had previously trained. After the data testing procedure, we also create a confusion matrix table to evaluate the model on the test data and give a sense of the SVM classification method's level of accuracy. In the framework of this study, evaluation measures are employed to assess how effectively the SVM model predicts the state of clean and healthy living behavior in coastal Surabaya City homes. The accuracy number, one of the performance indicators, will show an overview of the overall percentage of accurate predictions [31], and the recall value will guarantee that the model can identify the real CHLB [32]. The entire performance of the SVM model in this classification assignment will be fully depicted by the F1-score[32]. We assess the significance of the predictor variables last. This phase tries to comprehend how the predictor variables used in this research relate to one another and how these variables affect our categorization or prediction outcomes.

Table 1. List of Health Centers in the Coastal Area of Surabaya

Sub-Districts	Community Health Center	Sub-Districts	Community Health Center	Sub-Districts	Community Health Center
Asemrowo	Asemrowo	Krembangan	Krembangan Selatan	Rungkut	Kalirungkut
Benowo	Sememi		Dupak		Medokan Ayu
Pakal	Benowo		Morokrembangan	Gunung Anyar	Gunung Anyar
Pabean Cantikan	Perak Timur	Bulak	Kenjeran	Sukolilo	Menur
Semampir	Pegirian	Kenjeran	Tanah Kali Kedinding		Klampis Ngasem
	Sawah Pulo		Sidotopo Wetan		Keputih
	Sidotopo		Bulak Banteng	Mulyorejo	Mulyorejo
	Wonokusumo		Tambak Wedi		Kalijudan

3. RESULTS AND DISCUSSION

In this section, we will discuss the findings from our study on the SVM algorithm's usage to predict the CHLB status of coastal families in Surabaya. To provide readers with a better understanding of the findings, a thorough explanation will be included with these results.

3.1. Overview of Households with Clean and Healthy Living Behavior

Surabaya city has 12 areas which are included in the coastal area of Surabaya city, and there are 24 public health centers spread over the coastal area. Based on the data published by the Surabaya City Health Office, it is known that throughout the period 2016-2018 (Figure 4-6), there was a decrease in the number of households with the CHLB. According to the Puskesmas, this was indicated by the declining attainment of CHLB households in each health center. the target set by the Ministry of Health in the RPJMN 2015-2019, which is 80% of households with clean and healthy behaviors.

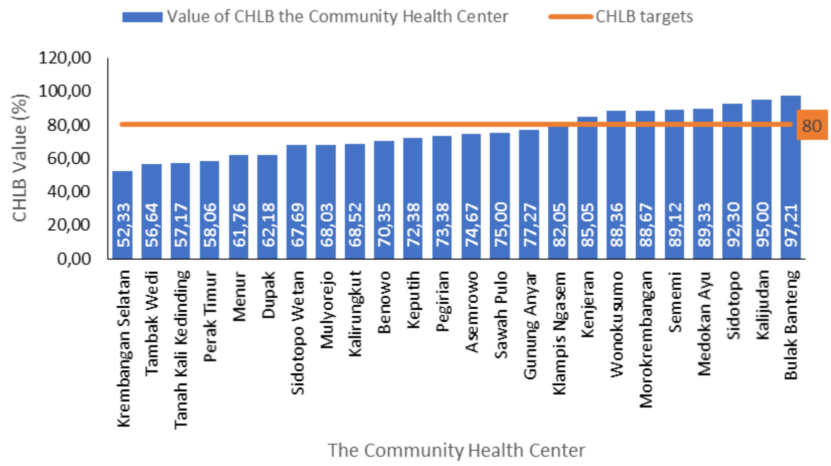


Figure 4. Overview of Households with CHLB by Health Center in 2016

In 2016 (Figure 4), there were nine health centers whose household achievements with CHLB exceeded the target of the RPJMN, these include Klampis Ngasem (Sukolilo District), Kenjeran (Bulak District), Morokrembangan (Krembangan Selatan District), Sememi (Benowo District), Medokan Ayu (Rungkut District), Kalijudan (Mulyorejo District), Bulak Banteng (Kenjeran District) and two health centers in Semampir District, namely the Wonokusumo and Sidotopo health centers.

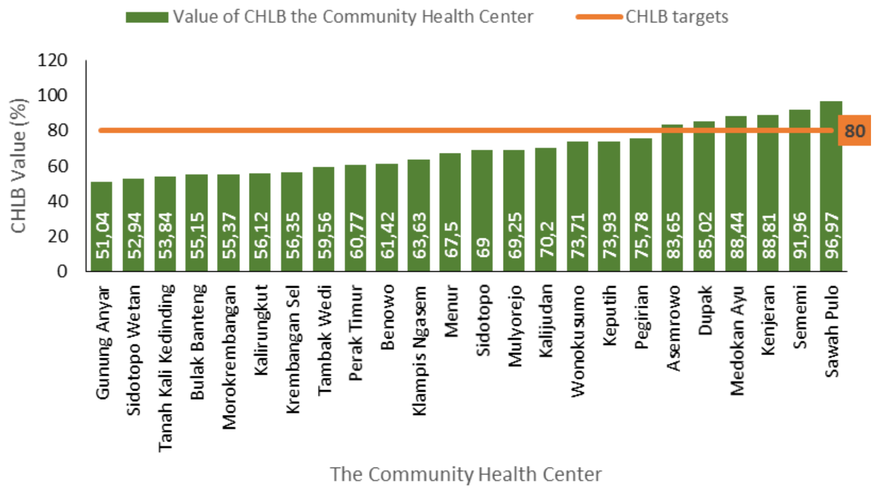


Figure 5. Overview of Households with CHLB by Health Center in 2017

Meanwhile, in 2017 (Figure 5), there was a decrease in the number of health centers that reached the CHLB target. There are only six health centers with CHLB achievements above 80 percent, including Asemrowo (Asemrowo District), Dupak (Krembangan Selatan District), Sawah Pulo (Semampir District), and three health centers that are still consistent from the previous year (in 2016) with household achievements with CHLB. above 80 percent, namely the Medokan Ayu Health

Center (Rungkut District), Kenjeran Health Center (Bulak District), and Sememi Health Center (Benowo District). Based on historical data in 2018 (Figure 6), there was another decline in the number of health centers that reached the CHLB target. There are only four health centers with CHLB achievements above 80 percent, including Pegirian Health Center (Semampir District), Dupak Health Center (South Krembangan District) which is still consistent from the previous year (in 2017), and three health centers that are still consistent from the previous two years (2016-2016). 2017) with the achievement of households having CHLB above 80 percent, namely the Medokan Ayu Health Center, and the Sememi Health Center (Benowo District).

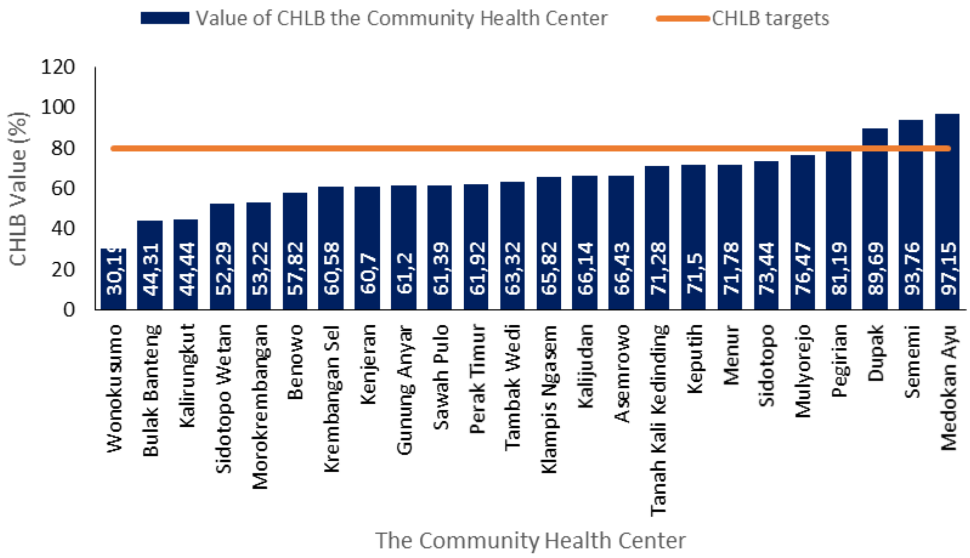


Figure 6. Overview of Households with CHLB by Health Center in 2018

3.2. Classification of Household Status with CHLB in the Coastal Area of Surabaya City

Using the SVM method by applying a linear kernel and the default parameter C, the value of parameter C is 1 [30], [33]. The following is the classification accuracy matrix, namely the confusion matrix as follows. Data on health facilities from the years 2016 to 2018 are used in this study, with a focus on those with low and good CHLB status. Based on the confusion matrix shown in Figure 7 and the training data, we discovered that 38 health centers had less CHLB status, and the SVM model was able to predict this status accurately. There are 16 research objects, though, for health facilities with a good CHLB classification. For one research object, the SVM model correctly predicted a good CHLB status, while it incorrectly predicted 15 other research objects.

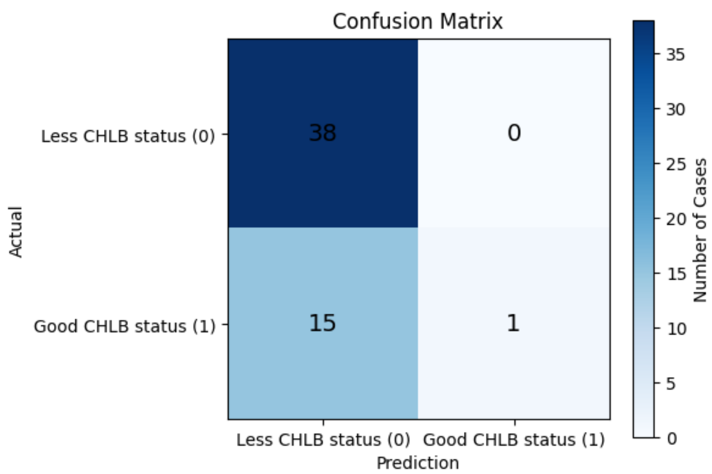


Figure 7. Confusion Matrix for the Surabaya coastline region's CHLB Status on the training dataset

Based on the confusion matrix shown in Figure 8 and the testing data. The test dataset is subsequently subjected to the same procedure. The goal of this is to assess model performance using unresearched data. Using the test data set, we were able to identify 15 community health centers between 2016 and 2018 with less CHLB status, and the prediction results confirmed this. The test data set did not predict the three community health centers that had good CHLB status from 2016 to 2018 to have bad CHLB status, even if this was false.

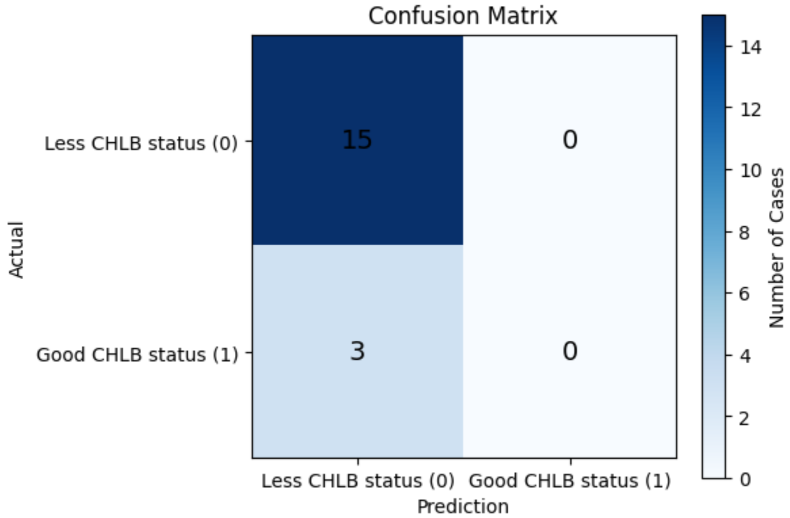


Figure 8. Confusion Matrix for the Surabaya coastline region's CHLB Status on the testing dataset

In the next stage, the model is evaluated using commonly established performance indicators. Table 3 summarises the conclusions of the evaluation. On the training dataset, the SVM model with a linear kernel achieved an accuracy score of 0.72. This means that the model correctly predicted 72% of the training data. The model accurately identified 72% of the positive cases in the dataset, due to a recall rate of 0.72 on the training dataset as well. The F1-Score for the training dataset is 0.84, indicating that precision and recall are well-balanced. This shows that the model did well in classifying the training set of data. Meanwhile, the model's performance on the test dataset improves dramatically. The SVM model has an accuracy rating of 0.83, which indicates it accurately predicted 83% of the test data. The recall rate on the testing dataset is equally high, at 0.83, indicating that the model is capable of identifying 83% of positive events in the testing dataset. What's more, the F1-Score value on the test dataset is 0.91, indicating that the model performs very well in predicting the CHLB status of homes in Surabaya's coastline area.

Table 2. Results of the CHLB Status Classification Evaluation Model

Dataset	Accuracy	Recall	F1-Score
Trainning Dataset	0.72	0.72	0.84
Testing Dataset	0.83	0.83	0.91

Testing data outperforms training data in terms of performance, which is an intriguing pattern in model evaluation outcomes. This implies that, in the context of this investigation, the SVM model with a linear kernel can produce more accurate predictions when provided with previously unseen test data. Gains in test data performance could result from a variety of factors. To begin, SVM models with linear kernels were probably successful in finding relevant patterns in training data that might be effectively generalized to test data. This illustrates that this methodology can discover relationships in data that are more complicated than would have been expected previously. Furthermore, greater performance on test data may imply that this model is not significantly overfitting to the training data. Overfitting occurs when a model is overly "specific" to the training data and cannot generalize successfully to fresh data. In this situation, the SVM model with a linear kernel appears to have a high amount of generalization. Finally, the improved performance on test data indicates that the SVM model with a linear kernel can offer reliable prediction results in practical circumstances for reliably predicting the CHLB status of homes in Surabaya City's coastal districts. The following step is to determine the importance of variables in the SVM model (Figure 9). This investigation uncovers crucial information about the elements that have the greatest influence on CHLB status in the Surabaya coastline area. The following is an interpretation of the variable importance level.

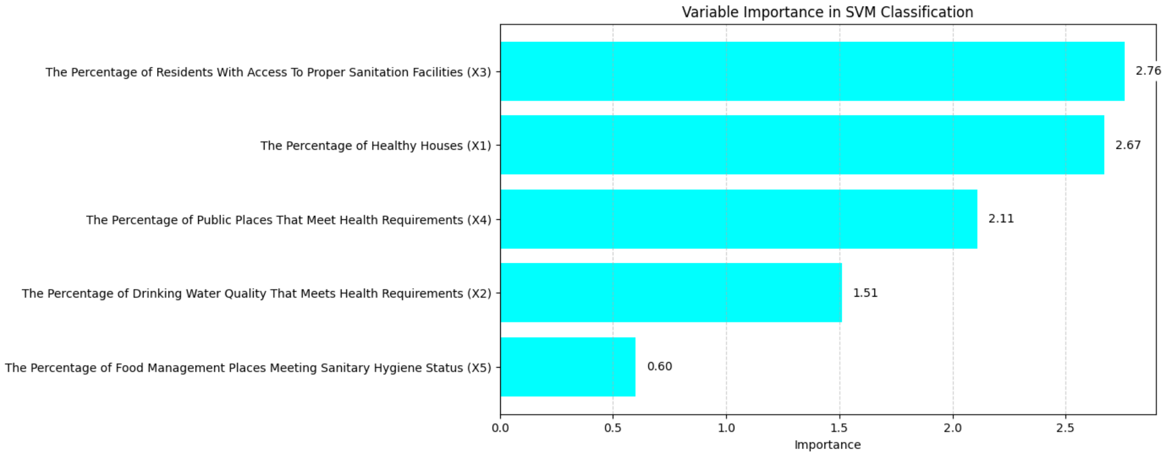


Figure 9. Variable Importance in SVM Classification for the Surabaya Coastline Region's CHLB Status

a. Percentage of the population with access to adequate sanitation facilities (X_3) (Importance Weight: 2.76)

This variable has the greatest importance and weight in the model, implying that it has a significant impact on the CHLB status of households in Surabaya's coastal area. The greater the proportion of the population with access to proper sanitary facilities, the more likely a household will have strong CHLB. This demonstrates that access to sanitary facilities is an important factor in determining clean and healthy living habits.

b. The percentage of residents with healthy homes (X_1) (Importance Weight: 2.67)

This variable has an elevated importance weight as well, showing that overall household conditions play a significant impact in determining CHLB. A clean, healthy, and well-maintained home is beneficial to CHLB.

c. Percentage of public facilities that meet health requirements (X_4) (Importance Weight: 2.11)

This variable demonstrates that the compliance of public locations with health norms has a considerable influence on home CHLB. The greater the number of public venues that satisfy health standards, the better the household CHLB in the area

d. Percentage of the population with access to quality drinking water that meets health requirements (X_2) (Importance Weight: 1.51)

The quality of drinking water is also crucial in CHLB. Households with access to safe drinking water have a lower risk of having CHLB.

e. Percentage of food management establishments that meet sanitation hygiene status (X_5) (Importance Weight: 0.60)

Even though it has a lower importance weight than other variables, the cleanliness of food processing facilities has an impact on household CHLB. Food processing facilities that fulfill hygienic requirements can help improve CHLB.

Thus, the findings of the analysis of the level of importance of these variables show that these variables play a significant role in the formation of CHLB in the Surabaya coastal area. This data can be utilized to develop more effective intervention programs and policies to increase CHLB in the region.

4. CONCLUSION

This study reveals significant findings concerning clean and healthy living behavior in Surabaya City's coastal districts. The accomplishment of CHLB targets in this region has deteriorated over time, according to a review of historical health data. This decline is a major source of worry in efforts to improve the quality of life in coastal communities. To address this issue, the study used machine learning methods, notably SVM with a linear kernel, to predict the CHLB status of homes in Surabaya's coastline area. The results of the investigation demonstrate that the SVM model can predict CHLB status with an accuracy of 83%. This is an encouraging sign that this technology can be used to monitor and improve CHLB in coastal areas. Furthermore, this study investigates the significance of several variables in impacting CHLB predictions. The percentage of the population with access to adequate sanitation facilities, the percentage of healthy homes, the percentage of public places that meet health requirements, the percentage of drinking water quality that meets health requirements, and the percentage of food management places that meet sanitary cleanliness status are all variables that have a significant impact on prediction. The findings of this study give significant information for policymakers and health practitioners to achieve the CHLB targets stated in the RPJMN. They can make more targeted and effective efforts by focusing on variables that have a substantial impact on improving CHLB in Surabaya City's coastal zones. Finally, this study contributes significantly to efforts to improve public health and quality of life in economically vital coastal districts of Surabaya.

REFERENCES

[1] S. Shirleyana, S. Hawken, and R. Y. Sunindijo, "City of Kampung: risk and resilience in the urban communities of Surabaya, Indonesia," *International Journal of Building Pathology and Adaptation*, vol. 36, no. 5, pp. 543–568, Oct. 2018, doi: 10.1108/IJBPA-02-2018-0025/FULL/XML.

[2] L. Sui, J. Wang, X. Yang, and Z. Wang, "Spatial-Temporal Characteristics of Coastline Changes in Indonesia from 1990 to 2018," *Sustainability 2020, Vol. 12, Page 3242*, vol. 12, no. 8, p. 3242, Apr. 2020, doi: 10.3390/SU12083242.

[3] M. Ishida, "Special Economic Zones and Economic Corridors," 2009.

[4] M. Shammi, M. M. Rahman, S. E. Bondad, and M. Bodrud-Doza, "Impacts of Salinity Intrusion in Community Health: A Review of Experiences on Drinking Water Sodium from Coastal Areas of Bangladesh," *Healthcare 2019, Vol. 7, Page 50*, vol. 7, no. 1, p. 50, Mar. 2019, doi: 10.3390/HEALTHCARE7010050.

[5] Y. Rizal, "Public response to the implementation of clean and healthy living behavior (PHBS) in coastal community in Rokan Hilir Regency," *Journal of Global Responsibility*, vol. 9, no. 3, pp. 261–279, Oct. 2018, doi: 10.1108/JGR-12-2017-0059/FULL/XML.

[6] S. F. Ayu, D. Aulia, and M. Nauly, "Improving Clean and Healthy Living Behavior in the Poor Communities in Coaltal Area, Serdang Bedagai," pp. 219–225, Dec. 2016, doi: 10.2991/PHICO-16.2017.42.

[7] T. Susanto, L. Sulistyorini, E. W. Wuryaningsih, and S. Bahtiar, "School health promotion: A cross-sectional study on Clean and Healthy Living Program Behavior (CHLB) among Islamic Boarding Schools in Indonesia," *Int J Nurs Sci*, vol. 3, no. 3, pp. 291–298, Sep. 2016, doi: 10.1016/J.IJNSS.2016.08.007.

[8] D. Susilaningrum and H. Al Azies, "Pemodelan Regresi Logistik Pada Faktor Yang Mempengaruhi Phbs Pada Rumah Tangga Penderita TBC Di Pesisir Surabaya," *EKSAKTA: Berkala Ilmiah Bidang MIPA*, vol. 18, no. 02, pp. 121–128, Nov. 2017, doi: 10.24036/EKSAKTA/VOL18-ISS02/65.

[9] H. Al Azies, "Analisis Perilaku Hidup Bersih Dan Sehat (PHBS) Rumah Tangga Penderita TB Di Wilayah Pesisir Kota Surabaya Menggunakan Pendekatan Regresi Logistik Biner," May 2017.

[10] Y. Susilowati and A. Santoso, "The Influence Of Our Schools Reducing Waste Program (Kurassaki) On Students 'Phbs In The Pilot Project School Of Bappeda, Tangerang District, Indonesia," *Proceedings of International Interdisciplinary Conference on Sustainable Development Goals (IICSDGs)*, vol. 3, no. 1, pp. 34–40, Dec. 2020, Accessed: Sep. 28, 2023. [Online]. Available: https://journals.ubmg.ac.id/index.php/IICSDGs/article/view/53

[11] H. Hasyim, Samsuryadi, M. E. Purnomo, B. B. Adhitya, N. A. Fajar, and H. Cahyono, "Community Empowerment To Improve Clean And Healthy Living Behavior [Chlb]: An Action Research," *International Journal Of Community Service*, vol. 1, no. 3, pp. 358–364, Nov. 2021, doi: 10.51601/IJCS.V1I3.41.

[12] R. Rachmathiany, B. M. Suhita, and N. Nurdina, "Factor that Influence on PHBS Students Female of Islamic Boarding School Al – Huda Kediri City Based on Health Belief Model Theory," *Journal for Quality in Public Health*, vol. 4, no. 1, pp. 265–270, Nov. 2020, doi: 10.30994/JQPH.V4I1.162.

[13] L. Khairiyati, F. Rahman, A. Udin, and V. Y. Anhar, "Factors affecting the implementation of clean and healthy livingbehavior at household level (Observational study at Sungai Paring village, Martapura Kota district)," *Indian J Public Health Res Dev*, vol. 10, no. 12, pp. 2037–2041, Dec. 2019, doi: 10.37506/V10/I12/2019/IJPHRD/192174.

[14] P. Mishra, C. M. Pandey, U. Singh, A. Gupta, C. Sahu, and A. Keshri, "Descriptive Statistics and Normality Tests for Statistical Data," *Ann Card Anaesth*, vol. 22, no. 1, p. 67, Jan. 2019, doi: 10.4103/ACA.ACA_157_18.

[15] J. Wexler, M. Pushkarna, T. Bolukbasi, M. Wattenberg, F. Viegas, and J. Wilson, "The what-if tool: Interactive probing of machine learning models," *IEEE Trans Vis Comput Graph*, vol. 26, no. 1, pp. 56–65, Jan. 2020, doi: 10.1109/TVCG.2019.2934619.

[16] C. Kim, R. W. Park, and S. J. Rhie, "174. Development of a Machine Learning Prediction Model to Select Empirical Antibiotics in Patients with Clinically Suspected Urinary Tract Infection using Urine Culture Data," *Open Forum Infect Dis*, vol. 8, no. Suppl 1, p. S196, Dec. 2021, doi: 10.1093/OFID/OFAB466.376.

[17] H. Kour, J. Manhas, and V. Sharma, "Comparative Study of Different Machine Learning Techniques in the Diagnosis of Dementia," *Advances in Intelligent Systems and Computing*, vol. 1187, pp. 93–104, 2021, doi: 10.1007/978-981-15-6014-9_12/COVER.

[18] N. M. Kebonye, "Exploring the novel support points-based split method on a soil dataset," *Measurement*, vol. 186, p. 110131, Dec. 2021, doi: 10.1016/J.MEASUREMENT.2021.110131.

[19] V. Vapnik, "The Support Vector Method of Function Estimation," *Nonlinear Modeling*, pp. 55–85, 1998, doi: 10.1007/978-1-4615-5703-6_3.

[20] C. U. Kumari, A. S. D. Murthy, B. L. Prasanna, M. P. P. Reddy, and A. K. Panigrahy, "An automated detection of heart arrhythmias using machine learning technique: SVM," *Mater Today Proc*, vol. 45, pp. 1393–1398, Jan. 2021, doi: 10.1016/J.MATPR.2020.07.088.

[21] L. Saulsberry *et al.*, "The social vulnerability metric (SVM) as a new tool for public health," *Health Serv Res*, vol. 58, no. 4, pp. 873–881, Aug. 2023, doi: 10.1111/1475-6773.14102.

[22] M. Awad and R. Khanna, "Support Vector Machines for Classification," *Efficient Learning Machines*, pp. 39–66, 2015, doi: 10.1007/978-1-4302-5990-9_3.

[23] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, Sep. 2020, doi: 10.1016/J.NEUCOM.2019.10.118.

[24] S. Ghosh, A. Dasgupta, and A. Swetapadma, "A study on support vector machine based linear and non-linear pattern classification," *Proceedings of the International Conference on Intelligent Sustainable Systems, ICISS 2019*, pp. 24–28, Feb. 2019, doi: 10.1109/ISS1.2019.8908018.

[25] P. C. Upadhyay, L. Karanam, J. A. Lory, and G. N. Desouza, "Classifying Cover Crop Residue from RGB Images: A Simple SVM versus a SVM Ensemble," *2021 IEEE Symposium Series on Computational Intelligence, SSCI 2021 - Proceedings*, 2021, doi: 10.1109/SSCI50451.2021.9660147.

[26] M. Saber, A. El Rharras, R. Saadane, A. H. Kharraz, and A. Chehri, "An optimized spectrum sensing implementation based on SVM, KNN and TREE algorithms," *Proceedings - 15th International Conference on Signal Image Technology and Internet Based Systems, SISITS 2019*, pp. 383–389, Nov. 2019, doi: 10.1109/SITIS.2019.00068.

[27] V. Sundararaghavan and N. Zabaras, "Classification and reconstruction of three-dimensional microstructures using support vector machines," *Comput Mater Sci*, vol. 32, no. 2, pp. 223–239, Feb. 2005, doi: 10.1016/J.COMMATSCI.2004.07.004.

[28] X. Tian, H. Daigle, and H. Jiang, "Feature Detection for Digital Images Using Machine Learning Algorithms and Image Processing," *SPE/AAPG/SEG Unconventional Resources Technology Conference 2018, URTC 2018*, Jul. 2018, doi: 10.15530/URTEC-2018-2886325.

[29] M. A. Almaiah *et al.*, "Performance Investigation of Principal Component Analysis for Intrusion Detection System Using Different Support Vector Machine Kernels," *Electronics 2022, Vol. 11, Page 3571*, vol. 11, no. 21, p. 3571, Nov. 2022, doi: 10.3390/ELECTRONICS11213571.

[30] H. Al Azies and G. Anuraga, "Classification of Underdeveloped Areas in Indonesia Using the SVM and k-NN Algorithms," *Jurnal ILMU DASAR*, vol. 22, no. 1, pp. 31–38, Jan. 2021, doi: 10.19184/JID.V22I1.16928.

[31] S. Gupta, K. Saluja, A. Goyal, A. Vajpayee, and V. Tiwari, "Comparing the performance of machine learning algorithms using estimated accuracy," *Measurement: Sensors*, vol. 24, p. 100432, Dec. 2022, doi: 10.1016/J.MEASEN.2022.100432.

[32] A. M. H. Taha, D. R. Syaiba Balqish Binti Ariffin, and S. S. Abu-Naser, “A Systematic Literature Review Of Deep And Machine Learning Algorithms In Brain Tumor And Meta-Analysis,” *J Theor Appl Inf Technol*, vol. 15, no. 1, 2023, Accessed: Sep. 29, 2023. [Online]. Available: www.jatit.org

[33] M. Abdar and V. Makarenkov, “CWV-BANN-SVM ensemble learning classifier for an accurate diagnosis of breast cancer,” *Measurement*, vol. 146, pp. 557–570, Nov. 2019, doi: 10.1016/J.MEASUREMENT.2019.05.022.