The Implementation of Support Vector Machine Method with Genetic Algorithm in Predicting Energy Consumption for Reinforced Concrete Buildings

Asep Syaputra
Teknik Informatika, Institut Teknologi Pagar Alam, Pagar Alam, Indonesia
Asepsyaputra68@itpa.ac.id

Abstract: Accurate information on energy consumption is crucial for measuring energy efficiency and savings in buildings. It refers to the energy needed to power a building at a specific time. Energy savings can reduce costs and environmental impact by lowering greenhouse gas emissions. Obtaining precise energy consumption data is essential for all parties involved in building planning, construction, and management. Over the past decades, global energy consumption in buildings has consistently increased, with HVAC systems being a significant contributor. To tackle this problem, research developed a support vector machine model with genetic algorithms to accurately predict energy consumption in buildings. Two models were tested: a standard support vector machine and a genetic algorithm-integrated support vector machine. The test results revealed that the support vector machine model achieved an RMSE value of 2.6. Additionally, the genetic algorithm optimized the parameter C and selected the most relevant predictor variables, reducing the RMSE to 1.7 and utilizing only 3 predictor variables. In the subsequent stage, parameter optimization and function selection were performed to achieve an improved RMSE value of 1.537. This research aims to enhance energy consumption prediction for reinforced concrete buildings by combining SVM and Genetic Algorithm. SVM serves as the primary prediction model, while the Genetic Algorithm is employed to determine optimal SVM parameters and relevant features. Recent studies have demonstrated that this combination yields more accurate predictions compared to standard methods. It enables more efficient energy planning, reduced operational costs, and optimized resource utilization in reinforced concrete buildings. However, it's worth noting that this implementation may require substantial processing and resource utilization, depending on the dataset's size and complexity.

Keywords: Genetic Algorithm, Support Vector Machine, Reinforced Concrete Building, Energy Consumption.

INTRODUCTION

Continuously increasing energy consumption has become a global issue that requires solutions to reduce energy consumption and greenhouse gas emissions. One way to address this is by constructing buildings that are more energy-efficient (Kumala et al., 2021). Based on the latest data published by the Indonesian National Standardization Agency (SNI) 1726 - 2012 regarding seismicity, since Indonesia is classified as a country prone to major earthquakes, it is indicated that the platform or extreme ground

*Asep Syaputra

motion conditions in many places in Indonesia can be classified as seismic design classes D, E, or F. These conditions can significantly impact engineering actions in civil works or other building structures and pose a high risk of collapse, endangering human life (Fauzan & Sapei, 2018).

To design earthquake-resistant high-rise structures, they must possess sufficient strength and perform well under various levels of loading, including the most crucial earthquake loads (Prasetyo & Bukhori, 2019). During an earthquake, forces corresponding to the stiffness of the structure come into play, and the structure responds until it collapses. When designing earthquake-resistant buildings, it is expected that the structure can adequately withstand seismic loads, ensuring that the building does not suffer damage during small to moderate earthquakes and does not collapse during major earthquakes (Prihatiningrum, 2020). The latest earthquake regulation is SNI 03-1726-2012, which aims to revise the previous regulation from 2012, namely SNI 03-1726-2002. This regulation takes into account the occurrence of seismic phenomena in Indonesia, which has resulted in substantial infrastructure damage. The seismic zone classification is divided into 6 seismic zones, providing more detailed information compared to the previous version. The spectral response and acceleration values are used to determine the level of seismic risk in different areas. Low-risk earthquake areas range from 0 to 0.15 g, moderate-risk zones range from 0.15 g to 0.5 g, and high-risk areas have acceleration values greater than 0.5 g (Masril, 2019).

SNI 03-1726-2012 stipulates that the structural system for resisting the horizontal, axial, and moment forces caused by an earthquake can be achieved through a moment-resisting frame system. This frame system comprises structural components and concrete energy dissipation required to provide strength to the building at a given time (Masril, 2019). Building energy consumption is becoming increasingly important due to growing concerns about inefficient energy usage and its negative impact on the environment. Therefore, predicting building energy consumption is a crucial step in energy savings that benefit individuals and society. By predicting energy consumption, we can design new buildings more wisely and enhance energy efficiency to save energy and reduce environmental impacts (Savitri, 2019). There have been many studies conducted to predict energy consumption in reinforced concrete structures with high accuracy using various computational methods and different datasets. These studies have adopted two types of calculation methods, namely individual variable methods and combined methods. In optimizing each problem, computers can be utilized if the variables involved can be transformed into electronic form. In this way, solutions to the problem can be easily obtained if the data related to the problem can be processed by a computer. Optimization can be performed to solve a problem by searching for the most advantageous conditions (Syaputra, 2022b).

According to the Asphalt Institute (1993), to achieve the required pavement quality, attention must be given to the presence of aggregate, which refers to the hard and rigid material used in the mixture. The requirements for asphalt concrete or asphalt concrete binder (AC-BC) mixtures include strength, durability, flexibility, shear strength, rutting resistance, ease of processing, water resistance, and cost-effectiveness. To assess the suitability of the aggregate used in asphalt mixtures, Marshall testing is conducted (No, 2019). In research conducted in China, the decision tree method was used to predict the level of concrete energy consumption. Meanwhile, in another study in China, Li, Meng, and their colleagues utilized Support Vector Machine to predict hourly cooling loads during summer in an office building. In Greece, Ekonomou conducted a study that employed Artificial Neural Network to predict long-term energy consumption (Dewi & Lumbanraja, 2017).

Support Vector Machine (SVM) is a useful algorithm for solving problems with small, non-linear, and multidimensional samples. SVM aims to minimize structural risk rather than empirical risk. SVM is used to map data from the input space nonlinearly to a multidimensional feature space, where the optimization of the classification model uses a linear hyperplane. When using SVM, there are two problems that need to be addressed: selecting optimal features for SVM and tuning the best parameters. The selection of appropriate features can affect the parameter fit, and conversely, irrelevant features will reduce classifier accuracy (Drajana, 2017). To improve the efficiency and accuracy of the classifier, some features need to be eliminated. Additionally, proper parameter tuning is also crucial in enhancing the accuracy of the SVM classifier. Key parameters in SVM should be adjusted to ensure optimal classification or regression accuracy (SB & Suparwito, 2022).

*Asep Syaputra

The Support Vector Machine (SVM) method has been widely used in various fields, including building energy consumption prediction. SVM can process high-dimensional data and perform well in regression problems. However, SVM has some drawbacks, such as sensitivity to parameters and high error rates. Genetic Algorithm (GA) is an optimization technique that can help improve the performance of SVM by optimizing SVM parameters (Husada & Paramita, 2021). To obtain more accurate predictions from the Support Vector Machine, parameter optimization such as the number of hidden layers and learning rate is necessary. Genetic Algorithm is used to efficiently search for these parameters. The fundamental principle of genetic algorithm is to search for better offspring through a selection process similar to natural selection, with the main goal of finding the optimal values of SVM model parameters. Genetic Algorithm is highly suitable for optimizing manifold and multi-objective parameters in Marshall testing by determining the optimal structure and parameters for the SVM model (Huda & Kom, 2019).

LITERATURE REVIEW

Feature selection is used to identify a subset of strong predictors within a database and reduce the number of predictors used in the calculation process (Syaputra, 2022a). This affects various aspects of the classification model, including the accuracy of the trained classification algorithm, the time required to train the classifier, the number of instances needed to train the model, and the costs associated with the function. Sometimes not all predictors are equally important in a specific application, and better efficiency can be achieved by removing some irrelevant or disruptive predictors. Therefore, we perform feature selection by eliminating useless data such as noise, outliers, and redundancies while maintaining the integrity of relevant and useful data. Data mining is used as a process to address problems by analyzing existing data in a database. Data mining is defined as the process of discovering patterns in data. Data mining, also known as Knowledge Discovery In Database (KDD), is the activity of collecting and utilizing historical data to uncover regularities, patterns, or relationships within a large dataset.

Support Vector Machine (Svm)

In 1992, Vapnik, along with his colleagues Bernhard Boser and Isabelle Guyon, introduced the Support Vector Machine (SVM). The SVM algorithm can handle high-dimensional training data by using non-linear mapping, allowing it to find a hyperplane for linear decomposition. In this technique, SVM utilizes support vectors and margin to search for the best classifier function that can separate two different classes. The goal of SVM is to find the optimal separating function that can distinguish between two types of objects, and the best hyperplane is the one that lies in the middle of the object clusters from the two classes (Nurachim, 2019). This technique has been widely used in classification and regression, and SVM is a mathematical model that can separate two or more classes of data by using a hyperplane in the feature space.

SVM functions by searching for a hyperplane with the largest margin between the two classes, where the margin is calculated as the distance between the hyperplane and the closest data points from each class. SVM can also be used to address the issue of overfitting by introducing a parameter called "C," which controls the trade-off between the margin and classification errors. Additionally, SVM can handle regression problems by introducing a kernel function, which can transform the data into a higher-dimensional feature space (Jumeilah, 2017). Kernel functions can help SVM discover more complex relationships between data.

Some advantages of SVM are (Jumeilah, 2017):

1. SVM can handle high-dimensional data effectively.
2. It is effective in cases where the number of features is larger than the number of samples.
3. SVM can handle non-linear data by using kernel functions.

*Asep Syaputra

4. It has a solid theoretical foundation and provides a good generalization performance.
5. SVM is less prone to overfitting compared to other classification algorithms.
6. It can handle both binary and multi-class classification problems.
7. SVM can work well with small sample sizes.
8. It provides control over the trade-off between margin and classification error through the regularization parameter.

SVM finds the most suitable hyperplane by maximizing the distance or margin between two different sets of classes. In SVM for linear classification, the optimization problem can be formulated as follows:

$$\min \frac{1}{2} \|\omega\|^2 \qquad (1)$$

Subject to

$$y_i(wx_i+b) \geq 1, \; i=1,\ldots,\lambda \qquad (2)$$

The common steps to perform classification using the SVM method are as follows:
1. Collect and preprocess the training data.
2. Select the appropriate SVM algorithm and kernel function.
3. Determine the hyperparameters of the SVM model, such as the regularization parameter and kernel parameters.
4. Train the SVM model using the labeled training data.
5. Evaluate the performance of the trained model using validation data or cross-validation.
6. Fine-tune the model by adjusting the hyperparameters if necessary.
7. Once the model is deemed satisfactory, use it to classify new unseen data.
8. Monitor and assess the performance of the SVM model in real-world scenarios.
9. Update and retrain the model periodically if new data becomes available or the classification requirements change.

GENETIC ALGORITHM

Genetic algorithms are optimization algorithms inspired by the principles of evolution in nature. The main principles of genetic algorithms are natural selection and genetic recombination, where individuals that are most suitable or fit for the environment pass on their genes to the next generation. In the context of predicting energy consumption for reinforced concrete buildings, genetic algorithms are used to search for the best parameters in the energy consumption prediction model using the Support Vector Machine (SVM) method (Christioko et al., 2022).

The optimized parameters through genetic algorithms are the C parameter and the gamma parameter in the Radial Basis Function (RBF) kernel of SVM. The C parameter controls the tolerance level for errors in the model, while the gamma parameter controls the width of the kernel function. The parameter optimization process starts by creating an initial population consisting of a number of random candidate solutions. Each candidate solution is scored based on its accuracy in predicting energy consumption on the training data. Then, the candidate solution with the best score is derived as a new solution for the next generation through genetic recombination and mutation. This process is repeated until reaching a predetermined termination condition, such as a certain number of iterations or a specific fitness value (Saukani et al., 2016).

By using a genetic algorithm on SVM, the energy consumption prediction model for reinforced concrete buildings can achieve better accuracy compared to SVM without the genetic algorithm. Additionally, incorporating a genetic algorithm in SVM can accelerate the model training time and reduce overfitting. However, it should be noted that the parameter optimization process in the genetic algorithm may require a considerable amount of time and significant computational resources. Therefore, selecting the appropriate parameters is crucial to optimize the utilization of the genetic algorithm in predicting energy consumption for reinforced concrete buildings.

METHOD

Research Design

The Research Design for Implementing Support Vector Machine Method with Genetic Algorithm in Predicting Energy Consumption for Reinforced Concrete Buildings can be designed with the following steps:

*Asep Syaputra

Fig. 1 Research Design

1. Research Objective: The objective of this research is to test the effectiveness of implementing the Support Vector Machine method with Genetic Algorithm in predicting energy consumption for reinforced concrete buildings.
2. Data: The data used in this research is energy consumption data of reinforced concrete buildings collected from several buildings in a specific area. The data should be complete with input variables such as building area, building usage type, usage time, weather conditions, and output variables in the form of energy consumption.
3. Method: The method used in this research is the Support Vector Machine method with Genetic Algorithm. The modeling process will be conducted by dividing the data into two parts: training data and testing data. The training data is used to build the prediction model, while the testing data is used to evaluate the accuracy of the constructed model.
4. Optimization Process: The optimization process is performed using the genetic algorithm to find the best parameters in the energy consumption prediction model. The parameters optimized through the genetic algorithm are C parameter and gamma parameter in the RBF kernel of SVM. The parameter optimization process is conducted on the training data.
5. Result Analysis: The result analysis is conducted by comparing the accuracy of the model constructed using the Support Vector Machine method with Genetic Algorithm to the accuracy of the model constructed without using the genetic algorithm. The model accuracy is evaluated using metrics such as Mean Squared Error (MSE) or R-squared.
6. Conclusion: The conclusion of this research demonstrates the effectiveness of implementing the Support Vector Machine method with Genetic Algorithm in predicting energy consumption for reinforced concrete buildings. The research findings can be used as a guide to improve energy efficiency in reinforced concrete buildings by optimizing energy usage.
7. Recommendations: The recommendations provided include further development of this research by considering other variables that can affect energy consumption in reinforced concrete buildings, such as building material type and building orientation. Additionally, the research can be conducted using a broader range of data and a diversified sample of reinforced concrete building types.

Data Collection

Data Collection can be done by gathering information from literature sources or by collecting data from previously constructed reinforced concrete buildings. The collected data includes the identified variables. Variables that have an impact on the energy consumption of reinforced concrete buildings are identified, such as building area, building orientation, wall materials, number of windows, cooling

*Asep Syaputra

systems, and so on. The data used in this study are secondary data obtained from the UCI Machine Learning Repository related to Energy Efficiency. There are 768 buildings of 12 different shapes, some of which have glass areas while others do not. The parameters used in this study include building energy performance factors, building construction, and analysis of thermal properties of physical materials, such as relative density, surface area, wall area, roof area, total building height, orientation, glass area, and cooling load.

Table 1. Building Energy Consumption Efficiency Data

No	Relative Density	Area Size	Wall Area	Roof Area	Number of Building Heights	Orientation	Glass Area	Cooling Load
1	0,9	514,5	294,1	110,2	7,0	2	0	21,3
2	0,9	514,5	294,1	110,2	7,0	3	0	21,3
3	0,9	514,5	294,1	110,2	7,0	4	0	21,3
4	0,9	514,5	294,1	110,2	7,0	5	0	21,3
5	0,9	514,5	294,1	110,2	7,0	2	0,1	26,47
6	0,9	514,5	294,1	110,2	7,0	3	0,1	26,37
7	0,9	514,5	294,1	110,2	7,0	4	0,1	26,44
8	0,9	514,5	294,1	110,2	7,0	5	0,1	26,29
9	0,9	514,5	294,1	110,2	7,0	2	0,1	25,95
10	0,9	514,5	294,1	110,2	7,0	3	0,1	25,63

Data Processing

Data that does not have a value in each dataset will be removed and not used, as part of an effort to obtain quality data. Several techniques that can be used to achieve this are as follows:

1. Data validation can indicate that the quality of the input data is unsatisfactory due to incompleteness, noise, or inconsistencies. To address these issues, identification, correction, and removal of outlier/noise data, inconsistent data, and incomplete data (missing values) can be performed.
2. Data integration and transformation are carried out to improve the accuracy and efficiency of the algorithms. The data used in this study consists of numerical values that are then transformed into the Rapidminer software.
3. Data dimensionality reduction and discretization are performed to obtain a dataset with fewer attributes and records while still providing significant information.
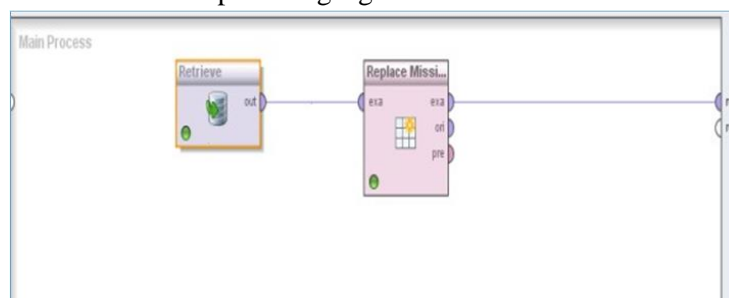

Fig. 2 Missing Data Replacement Scheme

Recommended Model

*Asep Syaputra

In the modeling phase, there are two types of data processed: training data (90%) and testing data (10%). In this stage, the algorithmic techniques to be tested will be described by inputting energy usage data and then analyzed and compared. The following is a description of the algorithmic techniques to be tested.
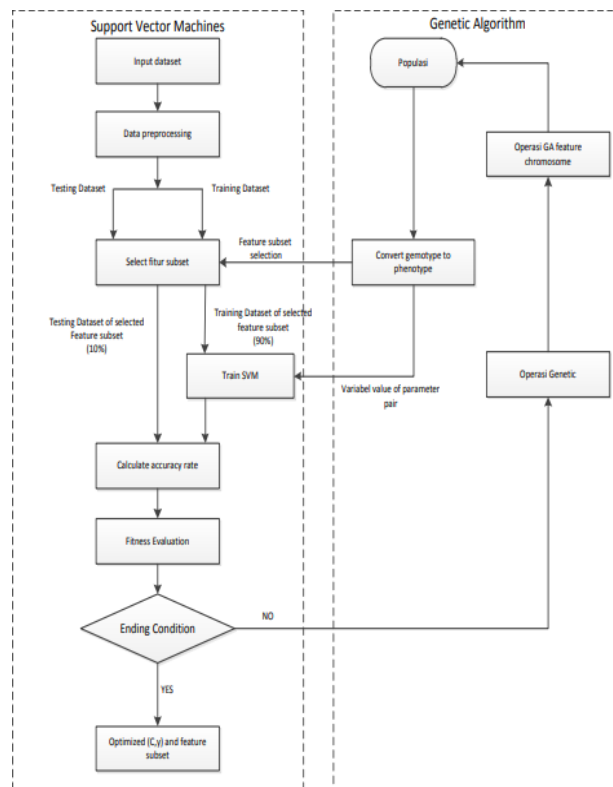

Fig. 3 The Recommended Model

Model Evaluation

This study consists of experimental stages and model testing using the energy efficiency dataset from the UCI Repository in the RapidMiner 5 application. In the experimental stage, specific software and hardware were used as aids, detailed in Table 2.

Table 2. Software and Hardware Parameters

Software	Hardware
Operating System Version: Microsoft Windows 7	Prosesor: Intel Pentium Dual Core
Data Processing: RapidMiner Version 5	Memory : 4 GB Hardisk : 500 GB

Testing And Validation

The output generated from this research is the validation process of the utilized model. Validation is conducted to evaluate the optimal predictive model by calculating the prediction error using the Root Mean Square Error (RMSE) value. This validation is performed by calculating the RMSE value, which is a commonly used evaluation metric for predictive models. RMSE measures the average difference between predicted values and actual observed values on the same scale. A lower RMSE value indicates better performance of the predictive model. The results of this validation process will indicate whether the used model is optimal or needs further improvement. If the validation results show a low RMSE value, then the model can be used to predict new data with a high level of accuracy. However, if the RMSE value is high, then the model needs to be improved or modified to enhance its performance.

*Asep Syaputra

RESULT

This research aims to validate the utilized model, where the validation is conducted with the purpose of evaluating the accuracy of the considered optimal predictive model, taking into account the Root Mean Square Error (RMSE) value as an indicator of prediction errors. By performing this validation process, it is expected to ensure that the model can provide accurate prediction results. In this regard, the model will be tested using relevant validation techniques to ensure that the results are reliable and valid. Additionally, the RMSE value will be used as a reference to measure the accuracy level of the predictive model. Thus, the results of this research are expected to contribute significantly to the development and improvement of better-quality predictive models in the future.

Results Of Testing And Implementation Of The Method

The results of testing and implementing a method are crucial parts of a research, particularly in the field of computer science and information technology. In this phase, researchers conduct trials and evaluations of the developed method or technique. The aim is to ensure that the method works effectively and efficiently, providing accurate and valid results. The results of testing and implementing the method are essential in determining whether the developed method can be widely used or not. Additionally, these results can provide valuable insights for researchers in improving and developing more effective methods in the future. Therefore, the testing and implementation phase is a critical stage in research and must be carried out meticulously and carefully.

Approach Using Svm Method

To evaluate the performance of the Support Vector Machine algorithm, testing is conducted using the K-Fold Cross Validation technique. This testing involves carefully selecting a subset of data for testing purposes to ensure that it represents the entire available data. The results of this testing are processed with the help of the RapidMiner application, facilitating the analysis and visualization of the obtained results. Thus, the use of the K-Fold Cross Validation technique and the RapidMiner application can assist in obtaining more accurate and efficient testing results.
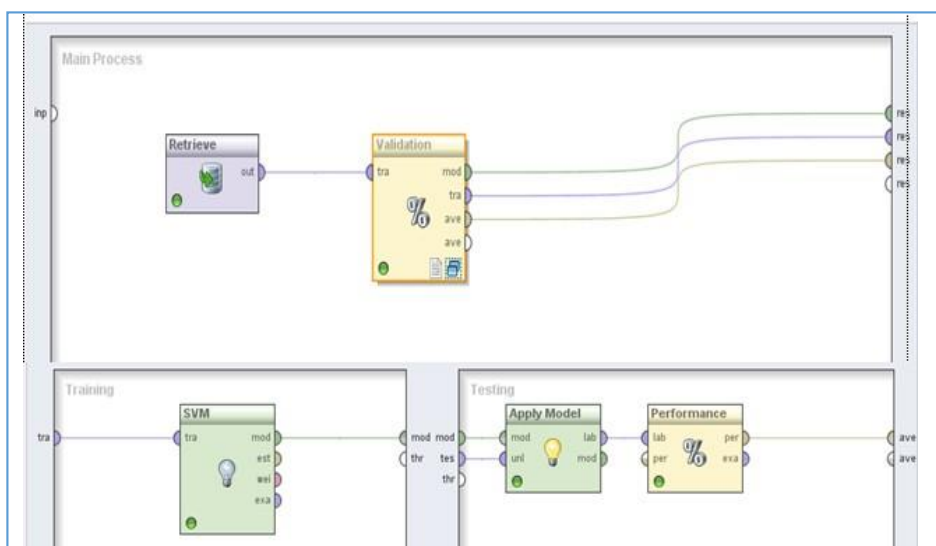

Fig 4. Evaluation using K-Fold Cross Validation in SVM Algorithm.

After undergoing a series of tests using the Support Vector Machine method, the evaluation results in the form of a Root Mean Square Error (RMSE) value of 2.613 were obtained. With the obtained test results, it can help improve the quality and accuracy of the prediction model used to support more precise and accurate decision-making.

Optimizing Svm Parameters With Genetic Algorithms

*Asep Syaputra

Below is an illustration of the testing results of the Support Vector Machine algorithm based on Genetic Algorithm for parameter optimization, performed using K-Fold Cross Validation technique and RapidMiner application. The K-Fold Cross Validation technique is used to test and evaluate the performance of the prediction model generated by the algorithm, while parameter optimization is conducted using genetic algorithms. The results of this testing can provide valuable information for improving the quality and accuracy of the prediction model.
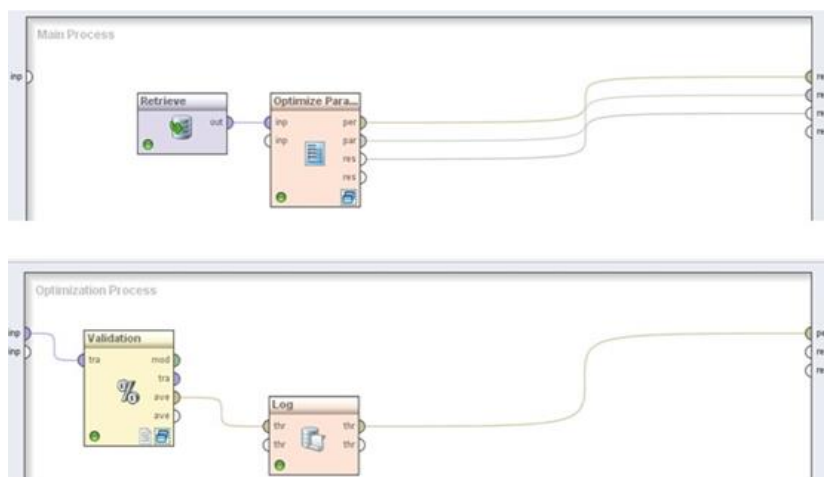

Fig 5. Testing of SVM Parameter Optimization Using Genetic Algorithm Approach

In order to enhance the performance and accuracy of the support vector machine model, an experiment was conducted to select the parameters C and γ using the Genetic Algorithm technique. The testing was carried out with the aim of obtaining the best parameters that can minimize the Root Mean Square Error (RMSE) value. After a series of tests, the evaluation result revealed an RMSE value of 1.825.

Feature Selection Method In Svm Using Genetic Algorithm

Here is an illustration of the testing of the Support Vector Machine algorithm based on the Genetic Algorithm technique for feature selection in the model. This testing utilizes the K-Fold Cross Validation technique to ensure the accuracy and reliability of the model. To perform this testing, the RapidMiner application is used as the main tool for data analysis and visualization of the experimental results.
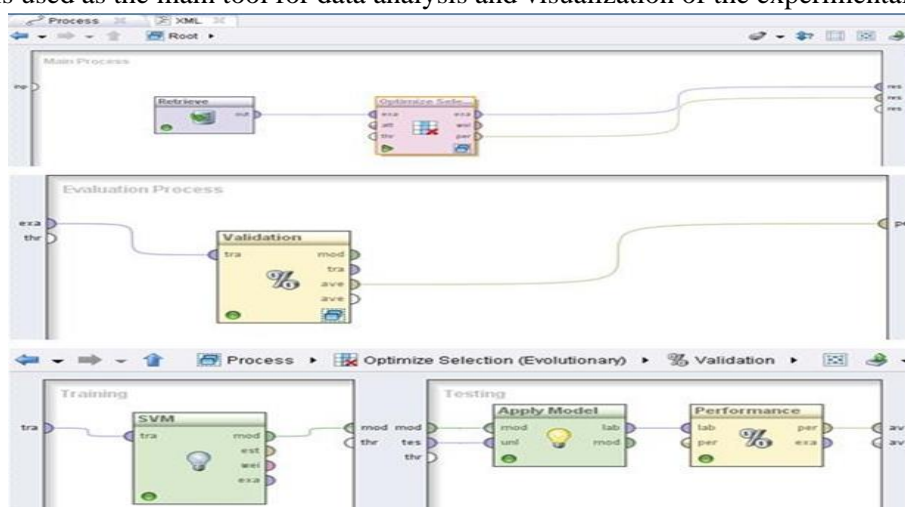

Figure 6. Testing Feature Selection of SVM using Genetic Algorithm Method

In this research, attribute or feature selection is conducted to determine the features that will be used in the prediction model. These features are selected based on their relevance in influencing the cooling load in buildings. The feature selection process considers the correlation values between each attribute

*Asep Syaputra

and the target variable to be predicted. Based on this analysis, the selected features include relative compactness, surface area, wall area, roof area, overall height, orientation, and glazing area. Additionally, one attribute is designated as the label, which is the cooling load that serves as the target prediction of the developed model.

Role	Name	Type	Statistics	Range	Missings
label	coolingload	real	avg = 24.296 +/- 9.648	[11.270 ; 46.230]	0
regular	relativecompactness	real	avg = 0.763 +/- 0.103	[0.620 ; 0.980]	0
regular	surfacearea	numeric	avg = 672.280 +/- 86.520	[514.500 ; 808.500]	0
regular	wallarea	numeric	avg = 317.275 +/- 41.842	[245.000 ; 416.500]	0
regular	roofarea	numeric	avg = 177.502 +/- 44.442	[110.250 ; 220.500]	0
regular	overallheight	numeric	avg = 5.250 +/- 1.759	[3.500 ; 7.000]	0
regular	orientation	integer	avg = 3.430 +/- 1.047	[2.000 ; 5.000]	0
regular	glazingarea	numeric	avg = 0.226 +/- 0.138	[0.000 ; 0.400]	0

Fig 7. Extraction of Best Features prior to Evaluation

Out of the 7 predictor variables, 3 attributes were selected that have an influence on the cooling load with an RMSE value of 1.767, namely wall area, overall height, and glazing area. The results of this attribute selection indicate that wall area, overall height, and glazing area are the most significant factors in influencing the cooling load in the tested system. Therefore, these three attributes can be the focus in the development of a more optimal and efficient system. Additionally, the use of this attribute selection technique can also help accelerate computation time and reduce the complexity of the used model.

Role	Name	Type	Statistics	Range	Missings
label	coolingload	real	avg = 24.588 +/- 9.513	[10.900 ; 48.030]	0
regular	wallarea	numeric	avg = 318.500 +/- 43.626	[245.000 ; 416.500]	0
regular	overallheight	numeric	avg = 5.250 +/- 1.751	[3.500 ; 7.000]	0
regular	glazingarea	numeric	avg = 0.234 +/- 0.133	[0.000 ; 0.400]	0

Fig 8. Optimal Feature Extraction after Testing

Feature Selection And Parameter Optimization Of Svm Using Genetic Algorithm

*Asep Syaputra

After performing feature selection and parameter optimization, the next step is to conduct testing using genetic algorithm to obtain the optimal RMSE value. In this testing, a Support Vector Machine algorithm based on Genetic Algorithm is used for feature selection and parameter optimization. During this testing process, multiple iterations are performed with different variations of parameters and features. The results of the testing are then analyzed to select the most optimal combination of features and parameters. By utilizing this technique, it is expected to discover a more accurate and efficient prediction model. The testing is conducted using RapidMiner application as shown in the following figure.
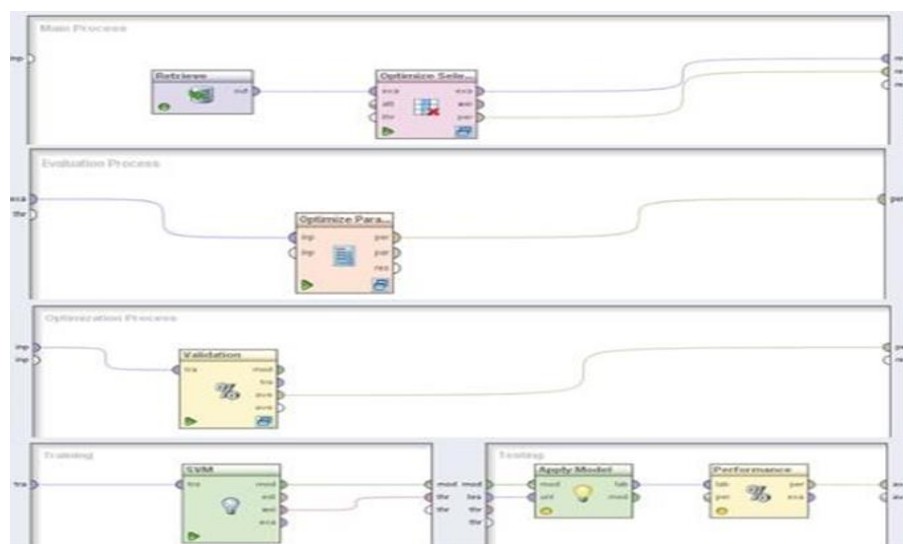

Fig 9. Testing Feature and Parameter Optimization of SVM using Genetic Algorithm Approach

The measurement result of RMSE of 1.537 indicates that the developed prediction model is capable of providing accurate estimates of cooling load. This demonstrates that the use of feature and parameter selection techniques with genetic algorithm in Support Vector Machine can improve the performance of the prediction model and achieve lower RMSE values compared to the prediction model without feature and parameter selection. In practical applications, the improved performance of the prediction model can bring significant benefits, such as cost and energy savings in cooling systems, as well as enhancing the quality and comfort of the indoor environment in buildings.

DISCUSSIONS

Cooling load plays a crucial role in predicting energy consumption in buildings. Based on the test results, an RMSE value of 2.613 indicates that the Support Vector Machine method can be utilized as an effective prediction tool for projecting energy consumption in buildings. This suggests that the factors influencing cooling load can serve as references in estimating energy consumption in buildings using the Support Vector Machine method. Although the RMSE value is still relatively high, the use of this method can be enhanced by adjusting the parameters and features employed to generate more accurate predictions.

The next step involves the utilization of a genetic algorithm to maximize the values of C and γ in order to optimize the RMSE value to 1.825. Additionally, the genetic algorithm is also employed for feature selection, resulting in an RMSE of 1.767. In the feature selection process, out of the 7 predictor variables, namely relative compactness, surface area, wall area, roof area, overall height, orientation, and glazing area, only 3 attributes, namely wall area, overall height, and glazing area, are selected as they influence the cooling load. Subsequently, parameter optimization and feature selection are performed to find the most optimal RMSE value, which is obtained as 1.537.
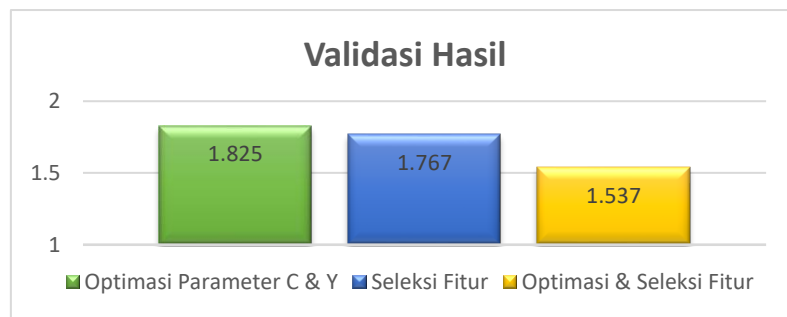
*Asep Syaputra

Fig 10. Validation Graph of Feature and Parameter Optimization Results for SVM

This research has significant implications for predicting the energy consumption of reinforced concrete buildings. These implications encompass several aspects, such as reducing operational costs, improving environmental comfort, and minimizing negative impacts on the environment. Additionally, the findings of this research can serve as a reference for the development of more efficient and sustainable energy technologies in buildings. Another implication is the importance of using support vector machine methods and genetic algorithms to enhance accuracy and effectiveness in predicting energy consumption in reinforced concrete buildings.

Based on the evaluation results, it is evident that the application of Genetic Algorithm to Support Vector Machines (SVM) for feature selection and parameter optimization can identify attributes that significantly influence energy consumption prediction with minimal error. Therefore, the Genetic Algorithm-based SVM method can be considered an effective approach for data classification. These findings demonstrate that the application of Genetic Algorithm to SVM can serve as a solution for predicting energy consumption in reinforced concrete buildings and assist architects in designing energy-efficient new structures. In Indonesia, this technique can be implemented by the Ministry of Public Works.

The use of Genetic Algorithm on Support Vector Machine (SVM) through RapidMiner software can contribute to decision-making in designing reinforced concrete buildings. This emphasizes the importance of managerial abilities of architects to plan the design of new structures more effectively. In the research context, the application of Genetic Algorithm to SVM can be used to evaluate the reliability of the method. Therefore, further research should involve the use of multiple datasets to enhance the validity of the results. Such research can be applied to other similar business units. Additionally, other methods such as Neural Network, Decision Tree, C4.5, and the like can be used for further development.

CONCLUSION

SVM based on Genetic Algorithm is used to perform attribute selection from the available 7 predictor variables. In this case, the use of Genetic Algorithm can help identify the most influential attributes in predicting energy consumption in buildings. From the experiments conducted on the energy efficiency UCI dataset, there is a significant difference between the average RMSE values before and after attribute selection and parameter optimization with Genetic Algorithm. This indicates that the use of SVM with Genetic Algorithm can improve accuracy in predicting energy consumption in reinforced concrete buildings with the smallest error value. The increased accuracy in predicting energy consumption in reinforced concrete buildings can assist architects and engineers in designing more efficient and environmentally friendly structures. Furthermore, this research can contribute to the development of technology and methods in the field of architecture and civil engineering, particularly in the use of SVM and Genetic Algorithm technology for data analysis in buildings and infrastructure. The results of this research can provide a better understanding of how data can be used to enhance the effectiveness and efficiency of energy use in buildings. Additionally, further research can explore other methods such as Neural Network, Decision Tree, C4.5, and similar approaches to further develop and expand the use of SVM and Genetic Algorithm technology in data analysis for buildings and infrastructure.

*Asep Syaputra
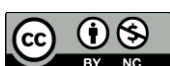
REFERENCES

Christioko, B. V., Asmiatun, S., & Susanto, S. (2022). Penjadwalan Kegiatan Praktikum Menggunakan Algoritma Genetika (Studi Kasus: Jurusan Teknologi Informasi Universitas Semarang). *Jurnal VOI (Voice Of Informatics)*, *11*(1), 25–34.

Dewi, D., & Lumbanraja, S. M. (2017). Rantai Pasok Industri Baja untuk Pembangunan PLTN di Indonesia. *Jurnal Pengembangan Energi Nuklir*, *19*(1), 51–60.

Drajana, I. C. R. (2017). Metode support vector machine dan forward selection prediksi pembayaran pembelian bahan baku kopra. *ILKOM Jurnal Ilmiah*, *9*(2), 116–123.

Fauzan, S. A., & Sapei, A. (2018). Evaluasi struktur gedung X di Jakarta berdasarkan SNI 03-1726-2012 ketahanan gempa untuk struktur gedung. *Jurnal Teknik Sipil Dan Lingkungan*, *3*(1), 11–24.

Huda, M., & Kom, M. (2019). *Algoritma Data Mining: Analisis Data Dengan Komputer*. bisakimia.

Husada, H. C., & Paramita, A. S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Teknika*, *10*(1), 18–26.

Jumeilah, F. S. (2017). Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *1*(1), 19–25.

Kumala, R., Ulpa, R., & Rahayu, A. (2021). Pajak Karbon: Perbaiki Ekonomi dan Solusi Lindungi Bumi. *Prosiding Seminar STIAMI*, *8*(1), 66–73.

Masril, M. (2019). Analisis Perilaku Struktur Atas Gedung Asrama Pusdiklat Ipdn Baso, Bangunan Wing 1 dengan Beban Gempa Berdasarkan Sni 03-1726-2012. *Rang Teknik Journal*, *2*(1).

No, A. I. (2019). *Teknologi Beton dan Bahan Bangunan*. Media Sahabat Cendekia.

Nurachim, R. I. (2019). Pemilihan model prediksi indeks harga saham yang dikembangkan berdasarkan algoritma Support Vector Machine (SVM) atau Multilayer Perceptron (MLP) studi kasus: saham PT Telekomunikasi Indonesia Tbk. *Jurnal Teknologi Informatika Dan Komputer*, *5*(1), 29–35.

Prasetyo, A., & Bukhori, B. (2019). Perbandingan Kebutuhan Tulangan Struktur Beton Bertulang Tahan Gempaterhadap Biaya Konstruksi Berdasarkan Pembebanan Sni 1726: 2012 Serta Sni 1727: 2013 Dengan Pembebanan Sebelumnya. *Logika Jurnal Ilmiah Lemlit Unswagati Cirebon*, *23*(1), 12–25.

Prihatiningrum, A. (2020). Karakter Bentuk Hunian Suku Rejang di Daerah Rawan Gempa (Studi Kasus: Desa Gunung Alam, Kabupaten Lebong). *ARSITEKTURA*, *18*(1), 84–93.

Saukani, M., Hakim, E. A., & Pakaya, I. (2016). Analisa Perbandingan Aliran Daya Optimal Mempertimbangkan Biaya Pembangkitan Dan Kestabilan Daya Menggunakan Particle Swarm Optimization Dan Algoritma Genetika. *Prosiding SENTRA (Seminar Teknologi Dan Rekayasa)*, *2*, 77–82.

Savitri, A. (2019). *Revolusi industri 4.0: mengubah tantangan menjadi peluang di era disrupsi 4.0*. Penerbit Genesis.

Sb, D. P. A., & Suparwito, H. (2022). Analisis Sentimen Bantuan Langsung Tunai Covid-19 Menggunakan Algoritma Support Vector Machine. *Prosiding Seminar Nasional Sanata Dharma Berbagi" Pengembangan, Penerapan Dan Pendidikan'sains Dan Teknologi'pasca Pandemi"*, 72.

Syaputra, A. (2022a). Analisis Kombinasi Metode Algoritma MFEP dan AHP Pada Pemilihan Bibit Unggul Kopi Robusta. *Komputika: Jurnal Sistem Komputer*, *11*(2), 185–192.

Syaputra, A. (2022b). Implementasi Metode Random Sampling Pada Animasi Motion Grapich Herbisida Dan Fungisida. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, *11*(2), 142–147.

*Asep Syaputra