Journal of Technology and Computer (JOTECHCOM)
Vol. 2, No. 3, August 2025, pp. 138~145
ISSN: 3048-0477 (Media Online)

RASKIN Recipient Classification Model Using LibSVM Based on Particle Swarm Optimization

Nuranisah[1], Subhan Hafiz Nanda Ginting[2]
[1,2]Universitas Battuta
[1]nuranisahasriel123@gmail.com, [2]subhanhafiz16@gmail.com

Article Info

Article history:

Received August 07, 2025
Revised August 07, 2025
Accepted August 08, 2025

Keywords:

LibSVM
RASKIN
Optimize
Data Mining
Discovery in Database

ABSTRACT

Receiving subsidized rice assistance (RASKIN) is a government program to distribute basic food assistance to underprivileged families. The distribution is carried out once every three months in accordance with the predetermined allocation. However, in practice, there are still often inaccurate targets in determining beneficiaries. This research aims to assist the village government in determining the eligibility of prospective RASKIN beneficiaries objectively and on target. The classification model used is Support Vector Machine (SVM) with the Library for Support Vector Machine (LibSVM) approach, combined with the Knowledge Discovery in Database (KDD) method. To improve classification performance, model parameter optimization is performed using the Particle Swarm Optimization (PSO) algorithm. Radial Basis Function (RBF) kernel is used in this process. The evaluation results show that the LibSVM model optimized with PSO is able to achieve an accuracy rate of 92.21%. The proposed model is expected to be an effective decision support system in selecting recipients of government social assistance more fairly and efficiently.

Corresponding Author:

Nuranisah
Universitas Batutta
Email: nuranisahasriel123@gmail.com

1. INTRODUCTION

Poverty is a problem caused by several interrelated factors such as income in the community that is still below the word feasible, unemployment, education, health, and transportation access to goods and services, geographical location, gender and environmental conditions [1]. Poverty is a source of social problems that have not been resolved in Indonesia properly so that many programs have been established to alleviate poverty. In July 1998 the Indonesian government launched a subsidized food assistance program to help the community called Operasi Pasar Khusus (OPK) Rice. This program continues to this day but the program is further expanded in its function to be right on target as to who is entitled to receive the subsidy assistance changed to For Poor Families (RASKIN) in 2002. [2] The benchmark for the success of the RASKIN plan is measured by several indicators of delivery, namely: data collection that is right on target, price, right amount, quality, time and accuracy of management. Each kelurahan has a big role in the process of smooth distribution of rice for poor families and must distribute it to people who are entitled to receive the benefits of the RASKIN program. [3]

The target of the poor rice subsidy is the middle to lower class community in Secanggang Village, Hulu and Hilir Hamlets. People who are entitled to receive RASKIN must have several criteria determined, the right data collection process is expected to help the subsidy process well and right on target. Machine learning is

considered to be one of the important things in the characteristics of artificial intelligence that supports the improvement of computerized systems that are able to pass on knowledge from past unprogrammed experiences in certain situations [4].

Data mining is a very important part towards the industrial revolution of era 4. Data mining cannot be separated from statistical analysis for classification methods. [5] Classification is the process of determining a model that explains the concept in a class of data whose purpose is to get the approximate results of an object whose label is not yet known [6]. The application of machine learning to the data management of RASKIN recipients is needed so that the analysis is carried out on target recipients. There are several algorithm writings used in the classification of RASKIN recipient data management, namely the C-4 algorithm, Random Forest, naïve Bayes, neural network, C4.5, k-nearest neighbor (k-NN) and support vector machine (SVM).

In previous research related to the classification of RASKIN recipients in several regions, algorithms have been carried out by writers such as analysis with the keywords "Raskin rice recipients" and "data mining" [7]. The research uses the C4.5 algorithm in classifying RASKIN data in Garogo District, in the classification results using the C4.5 algorithm can easily process data because of its easy adaptation to complex interactions between variables [8]. Using data in the analysis of RASKIN recipients has also been done with several Fuzzy K- Nearest Neighbor (FKNN) and Support Vector Machine classification methods using the SVM method better than FKNN with kernel parameters that function well worth C = 100 and□ = 0.000283 on SVM. [9]

Previously there was also raised research with the Support Vector Machine (LibSVM) method which proved good accuracy results for the classification of RASKIN recipient data in Wonosobo district with the best accuracy results on the Radial Basis Function (RBF) kernel function with a percentage of 83.1933%. [10]. In this study using Parameter estimation with the K-Fold Cross Validation method as a division between test data and training data.

Although various methods of analyzing RASKIN recipients have been carried out to find out the data collection that is right on target for people who are entitled to receive poor rice subsidy assistance (RASKIN), in this study what is done is to look at data collection with a classification approach that does not use cross validation with the Library Support Vector Machine (LibSVM) Model algorithm. With the background that has been listed, this research aims to analyze the acceptance of RASKIN using a classification approach with the Library Support Vector Machine (LibSVM) Model. The arrangement of the stages that will be carried out for this research is to collect data on the hamlet in the village and then make changes to several attributes to facilitate the calculation process that will be carried out, classifying the dataset using the Library Support Vector Machine (LibSVM) Model algorithm.

2. METHOD

There are 2 perspectives carried out in research, namely descriptive research where the approach process is qualitative. Qualitative descriptive method is a method that is processed using qualitative data and then described descriptively [11]. Qualitative data is descriptive data in the form of numerical symbols or qualitative numbers carried out so that it is easy to understand empirical phenomena, especially for as many descriptions as possible about these phenomena without having to go into detail in the relationship between variables that are interrelated with each other. [12] The process is carried out through several stages, namely data collection, initial classification, data preprocessing, attribute labeling to change text to nominal, data sharing, classification with SVM modeling (LibSVM) and evaluation and analysis. Figure 1 explains the process of the research flow that will be carried out

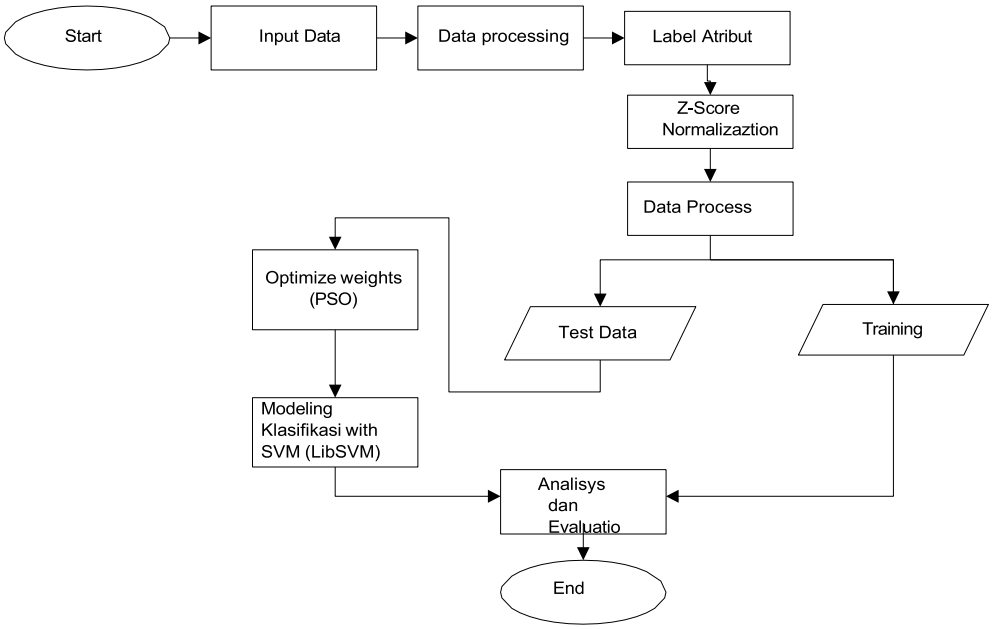

Figure 1. Illustration of support vector machine

a. Data Retrieval

This study classified data into 2 classes, namely "Yes" or "No". The population data used is data on the residents of Secanggang village, upstream and downstream hamlets, where there are 289 data on the heads of households of villagers who will be classified for RASKIN recipients.

b. Initial Classification

Before preprocessing, population data is classified manually first to make it easy to understand by making several changes such as sorting and converting some data into simple nominal data that has been grouped to make it easier to standardize the value later.

c. Data Preprocessing

At this stage, some standardization of values is carried out to make it easier to classify data, some samples in the dataset are converted from non-numeric format to numeric format. can be seen in table 1.

d. Attribute Labeling

The need for data preprocessing from machine learning, where the data will be represented to a numerical format so that it is easy to process the calculations that will be used in accordance with the application of the model. In the data preparation process, all activities are involved in the data set so that it can be applied to the classification model to be designed. [13]

Table 1. Table 1. Data conversion

No	Name of Head of Household	Name of Head of Household	Gender	Age	Status	Occupation	Number of dependents	Number of children	family members who, Income/ Monthly	Owned house (RMS Owned by parents/	Borrowed	Contract /Rent	Raski n
1	DUSUN HILIR	ABAS	1	48	1	6	2	1	0	4	1	0	0	0	Yes
2	DUSUN HILIR	ABD.HALID	2	61	1	9	2	1	0	4	1	0	0	0	Yes
3	DUSUN HULU DALAM	ABDUL KARIM	1	46	1	2	3	0	0	4	0	2	0	0	Yes
4	DUSUN HULU DALAM	ABDUL RAHMAN	1	55	1	6	4	0	3	5	1	0	0	0	Yes
5	DUSUN HULU DALAM	ABDUL WAHAB	1	48	1	7	3	2	0	4	1	0	0	0	No
6	DUSUN HILIR	ABDULLAH	1	41	1	6	5	4	0	5	1	0	0	0	Yes
7	DUSUN HILIR	ABDULLAH	1	58	2	6	0	0	0	1	1	0	0	0	No

At this stage the labeling is divided into 2 classes namely "Yes" and "No". In the process of determining data labeling, it is categorized based on the results of the initial data classification carried out to simplify calculations. Data labels are determined based on categories on several attributes that are used as a benchmark for the 2 classes.

2.2. Perform Optimize Weights (PSO)

Particle Swarm Optimization (PSO) is a method derived from natural patterns by groups of animals that live in groups such as termites, birds, ants or bees. The PSO algorithm is inspired by the characteristics of a natural trait of some of these organisms. The nature consists of habits that have been carried out in daily activities and influenced by individuals with other individuals in a population. In particle Swarn Optimization (PSO) there is a particle, the word "particle" is a reference to an individual, for example: a bird in a bird population [14]. One of the methods in PSO, Inertia Weight, which is each particle is an approach to the particle distance [15] Between one particle that is on the right or shorter path to the center point, the remaining particles will also follow that path even though the distance traveled is far.

2.3. SVM Classification Modeling (LibSVM)

Libsvm is an efficient SVM classification and package library that is useful for obtaining the optimal solution of duality language in SVM. The classification model has clear accuracy and can achieve optimal accuracy. [16] The vector classification in question is nu- SVC and c-support vector classification (C-SVC); epsilon-support vector regression (epsilon- SVR), nu-SVR for regression and distribution estimation in the form of one-class SVM [17].The use of the LibSVM program model follows 2 stages, namely by conducting a training stage on the dataset in order to obtain a model, then utilizing the model obtained so as to obtain predictive information in a training dataset

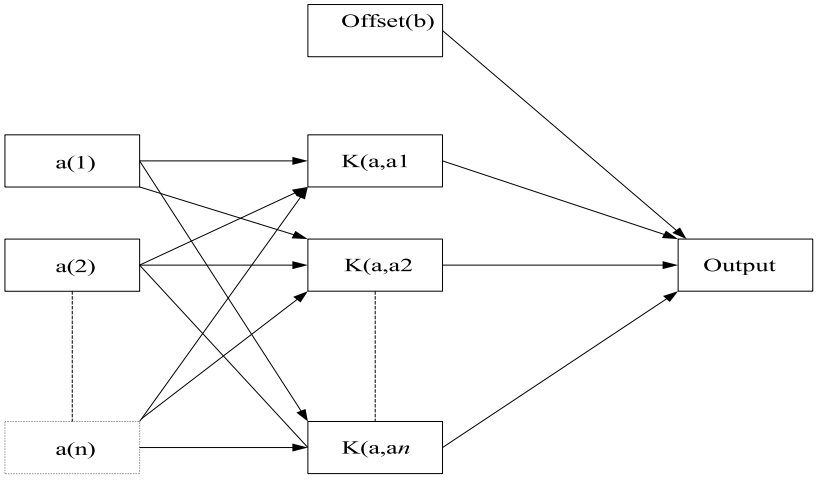

Figure 2. LibSVM Architecture

LibSVM is the completion of the previous SVM with the development of several parameters added. LibSVM functions to solve two-class problems by building Hyperlane and distinguishing two sides, namely "YES" and "NO". Figure 2 is a LibSVM architecture. In the research flow process in order to get the performance value of the model used in LibSVM, several kernel functions are applied. Some of the kernel functions used in this study are linear kernel (2), radial - basis function (rbf) kernel in (3), sigmoid kernel (4) and polynomial kernel (5).

$$K(x_i, x_j) = x_i^T, x_j \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$K(x_i, x_j) = \exp(-\frac{\|x_i, x_j\|^2}{2\sigma^2}) \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

$$K(x_i, x_j) = \tanh(\beta_0\, x_i^T, x_j + \beta_1) \ldots\ldots\ldots\ldots (4)$$

$$K(x_i, x_j) = (1 + x_i^T, x_j)^p \ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Where K is the class number, $K(xi, xj)$ is the kernel in the element xi, xj, $\beta 0$ is the classification that affects the out-comes, and$p$ is the degree set for the flexibility of the classifier.

2.4. Classification Model Analysis and Evaluation

At this stage the evaluation process is carried out in order to produce a performance value of the model built. The performance of the algorithm is measured using confusion matrix and measuring the value of accuracy, precision, recall and support vector.

2.5. Confusion Matrix

Classification determines 4 predictions generated, including confusion matrix, accuracy, precision and recall. There are 4 terms used in the method, namely:

1) True Positive (TP) : Prediction of the result of the amount of output data and the actual class of output
Positive

2) True Negative (TN) : Predicted number of data in the actual class is positive, but the projected class is negative.
is projected to be negative.

3) False Positive (FP) : The projected class is positive but the actual class
has a positive amount of data.

4) False Negative(FN) : The actual and predicted classes have a negative amount of data.

Table 2. Confusion Matrix

Prediction			
Fact	Positive	Negative	Neutral
Positive	TP	$FNg1$	$FNt1$
Negative	FP1	TNg	$FNt2$
Neutral	FP2	$FNg2$	TNt

The calculations for the results of *accuracy, precision* and *recall* are shown as follows:

$$Akurasi = \frac{(TP+TNt+TNg)}{(TP+FNg1+FNt1+FP1+TNg+FNt2+FP2+FNg2+Tnt)'} \ldots\ldots\ldots\ldots (6)$$

$$Presisi = \frac{TP}{TP+FP1+FP2'} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

$$Recall = \frac{TP}{TP+FNg1+FNt1'} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

Equation (6) accuracy shows the ratio of the number of correct predictions to the results of the number of input samples. Precision (7) is how many true positive values among all predicted positive values. recall in equation (8) shows how many positive predicted values.

3. RESULTS AND DISCUSSION

3.1. Data Collection

In the data collection process under household data in the Hulu and Hilir hamlets of Secanggang village, 298 data were collected with 12 regular attributes and 1 special attribute. Figure 3 shows the amount of data that received RASKIN assistance in the hamlet. A significant comparison of the number of recipients of RASKIN assistance requires normalization and using the Optimize Weights (PSO) method as data balancing for the next process.

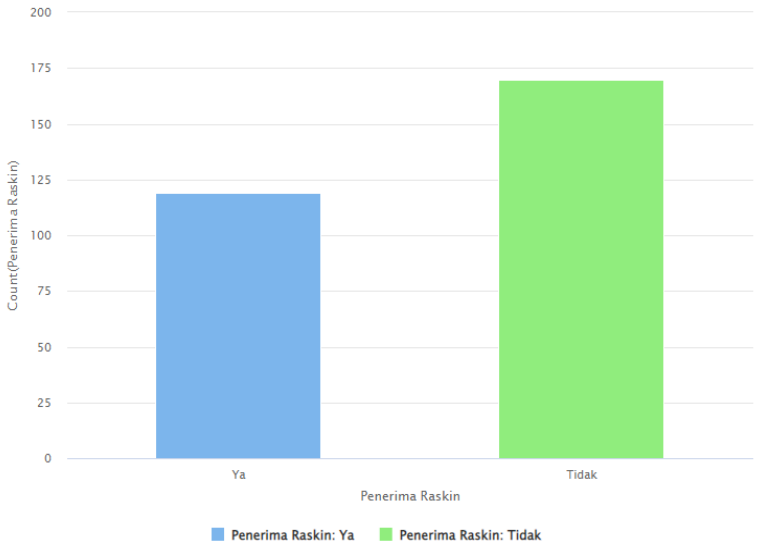

Figure 3. Comparison of the number of RASKIN recipients in Hulu and Hilir Hamlets

3.2. Data Division

In order to obtain the results of the classification of RASKIN beneficiaries and non-recipients of RASKIN assistance for households in the HULU and HILIR hamlets of Secanggang village, the following stages of analysis were carried out:

1. This study uses 298 data divided into training data and testing data with a ratio of 80: 20 in the KDD process only 12 attributes are used.
2. The use of kernel functions in this study are linear, Polynomial, Sigmod and Radial Basic Function (RBF) kernel functions.
3. To determine the best parameters, the Optimize Weights (PSO) method was used, which was first normalized with Z-score as a measure of the performance of the LibSVM model.

3.3. Classification Results of LibSVM method with Optimize Weights (PSO)

The comparative results of experiments using the LibSVM model there are several different levels of accuracy can be seen in table 3.

Table 3. Testing Results of LibSVM Method with Optimize Weights (PSO)

SVM Type	Kernel Function	Stratified Sampling	Shuffled Sampling	Linear Sampling
C-SVC	Rbf	76.19%	75.76%	72.73%
C-SVC	Polynomial	88.74%	86.58%	80.95%
C-SVC	Linear	58.87%	63.64%	59.74%
C-SVC	Sigmoid	54.11%	59.74%	62.77%
nu-SVC	Rbf	**92.21%**	91.34%	91.34%
nu-SVC	Polynomial	86.58%	89.18%	86.15%
nu-SVC	Linear	58.87%	60.17%	58.44%

nu-SVC	Sigmoid	57.58%	61.04%	58.87%
one-class	Rbf	54.11%	56.28%	50.65%
one-class	Polynomial	64.94%	61.90%	59.31%
one-class	Linear	58.87%	58.01%	57.58%
one-class	Sigmoid	61.47%	61.47%	55.41%

From the use of kernel type parameters in experimental results have different characteristics, based on the dataset used so that the performance results become the same and/or different, according to the types of kernels to be used with different datasets. Some of the kernel type parameters that have been shown from the results of the model performance values have a slight difference from each kernel, which is caused by the characteristics of the data used when preprocessing the data.

The results of the trial process are the stages of the experimentation process of the LibSVM model. The stage process is carried out by applying Optimize Weights (PSO) and the LibSVM model produces good accuracy of 92.21% by using RBF parameters in the stratified sampling method with the model validation process. Although it is good, in trials with linear sampling parameters with a one-class model the trial results obtained are very low at 50.65%.

Table 4. Accuracy Results based on Confusion Matrix

	True Negative	True Positive
Negative Prediction	85	8
Positive Prediction	10	128

Table 4 shows the comparison of the accuracy results of the RASKIN acceptance classification model using the LibSVM and Optimize Weights (PSO) models, showing that there are different results from each parameter applied using the LibSVM and Optimize Weights (PSO) models.

4. CONCLUSION

This study has classified the data analysis of RASKIN receipts in Hulu and Hilir hamlets of Secanggang village using LibSVM and Optimize Weights (PSO). The LibSVM and Optimize Weights (PSO) classification produces the greatest accuracy value of 92.21% by using the Stratified Sampling parameter in the RBF kernel function with a precision of 92.75% and a recall of 94.12%. From the research results it can be seen that the classification process using LibSVM and Optimize Weights (PSO) produces the best accuracy in the RBF kernel function.

REFERENCES

[1] D. Chauhan, Shivani, and P. N. Suganthan, "Learning strategies for particle swarm optimizer: A critical review and performance analysis," *Swarm Evol Comput*, vol. 98, p. 102048, Oct. 2025, doi: 10.1016/J.SWEVO.2025.102048.

[2] Y. Wang, Z. Liu, and H. Han, "Collaborative knowledge transfer-based multiobjective multitask particle swarm optimization," *Swarm Evol Comput*, vol. 98, p. 102115, Oct. 2025, doi: 10.1016/J.SWEVO.2025.102115.

[3] L. Qiu, Y. Zhu, C. Xu, G. Ren, Y. Hu, and X. Liu, "Physics-Informed Neural Network Optimized by Particle Swarm Algorithm for Accurate Prediction of Blast-Induced Peak Particle Velocity," *Intelligent Geoengineering*, Aug. 2025, doi: 10.1016/J.IGE.2025.07.002.

[4] P. Nie, Y. Gan, and Q. Yu, "Design an enhanced granulation-degranulation mechanism using group-exchange particle swarm optimization," *J Ind Inf Integr*, vol. 47, p. 100919, Sep. 2025, doi: 10.1016/J.JII.2025.100919.

[5] N. A. Tsedura, E. Bhero, and C. Chibaya, "Towards the design of a particle swarm optimization ontology for object classification," *Array*, vol. 27, p. 100449, Sep. 2025, doi: 10.1016/J.ARRAY.2025.100449.

[6] G. Zhou, Z. Fang, Y. Hu, J. Chen, and J. Ren, "An adaptive strategy quantum particle swarm optimization method based on intuitionistic fuzzy entropy and evolutionary game theory," *Appl Soft Comput*, vol. 183, p. 113654, Nov. 2025, doi: 10.1016/J.ASOC.2025.113654.

[7] L. Deng and S. Liu, "Collective dynamics of particle swarm optimization: A network science perspective," *Physica A: Statistical Mechanics and its Applications*, vol. 675, p. 130778, Oct. 2025, doi: 10.1016/J.PHYSA.2025.130778.

[8] O. O. Petinrin, X. Li, and K. C. Wong, "Particle Swarm Optimized Gaussian Process Classifier for Treatment Discontinuation Prediction in Multicohort Metastatic Castration-Resistant Prostate Cancer Patients," *IEEE J Biomed Health Inform*, vol. 26, no. 3, pp. 1309–1317, Mar. 2022, doi: 10.1109/JBHI.2021.3103989.

[9] A. Houkan *et al.*, "Enhancing Security in Industrial IoT Networks: Machine Learning Solutions for Feature Selection and Reduction," *IEEE Access*, vol. 12, pp. 160864–160883, 2024, doi: 10.1109/ACCESS.2024.3481459.

[10] S. Xu, R. C. Deo, O. Faust, P. D. Barua, J. Soar, and R. Acharya, "Automated Lightweight Model for Asthma Detection Using Respiratory and Cough Sound Signals," *Diagnostics 2025, Vol. 15, Page 1155*, vol. 15, no. 9, p. 1155, May 2025, doi: 10.3390/DIAGNOSTICS15091155.

[11] H. Hu, M. Li, and H. Zhang, "Dynamic surrogate-assisted particle swarm optimization for chemical kinetic mechanisms," *Comput Chem Eng*, vol. 202, p. 109294, Nov. 2025, doi: 10.1016/J.COMPCHEMENG.2025.109294.

[12] X. Luo and Z. Yi, "Efficiency management of engineering projects based on particle swarm multi objective optimization algorithm," *Systems and Soft Computing*, vol. 7, p. 200320, Dec. 2025, doi: 10.1016/J.SASC.2025.200320.

[13] M. Mohammadpour and A. Mohammadpour, "Cross-Material Forecasting-Optimization of CO2 adsorption using deep transfer learning and Particle Swarm Optimization: From porous carbons to Metal-Organic frameworks," *Sep Purif Technol*, p. 134655, Aug. 2025, doi: 10.1016/J.SEPPUR.2025.134655.

[14] M. Wang, W. Zheng, S. Zhao, and C. He, "Voronoi node stepwise following particle swarm optimization for path planning of underwater gravity aided navigation," *Ocean Engineering*, vol. 336, p. 121767, Sep. 2025, doi: 10.1016/J.OCEANENG.2025.121767.

[15] T. Varghese, N. K. Arun, and G. Pathirikkat, "Congestion management in power grids using multi-agent systems and particle swarm optimization," *Franklin Open*, vol. 12, p. 100303, Sep. 2025, doi: 10.1016/J.FRAOPE.2025.100303.

[16] X. Lu, M. Chen, and Y. Tian, "Accurate state-of-charge estimation of LiFePO4 battery: An adaptive extended kalman filter approach using particle swarm optimization," *Energy Reports*, vol. 14, pp. 1169–1178, Dec. 2025, doi: 10.1016/J.EGYR.2025.07.023.

[17] I. Bala, W. Karunarathne, and L. Mitchell, "Optimizing Feature Selection by Enhancing Particle Swarm Optimization with Orthogonal Initialization and Crossover Operator," *Computers, Materials and Continua*, vol. 84, no. 1, pp. 727–744, Jun. 2025, doi: 10.32604/CMC.2025.065706.