THE JOURNAL OF INDONESIA SUSTAINABLE DEVELOPMENT PLANNING

VOL. 5 NO. 3 - December 2024

E-ISSN: 2722-0842 | P-ISSN: 2721-8309

Available online at
http://journal.pusbindiklatren.bappenas.go.id/

Ministry of National Development Planning/Bappenas Republic of Indonesia

Research Paper

AI-Based Models for Identifying Underdeveloped Villages in Indonesia's Rural Development

Harun Al Azies

Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang, Indonesia

****) Correspondence author***: harun.alazies@dsn.dinus.ac.id

Abstract

Rural underdevelopment remains a significant barrier to achieving the Sustainable Development Goals (SDGs), particularly SDG 1 (No Poverty), SDG 10 (Reduced Inequality), and SDG 11 (Sustainable Communities), in Indonesia. This study addresses this challenge by employing Artificial Intelligence (AI) and machine learning techniques to predict and classify underdeveloped villages, offering a data-driven approach to inform targeted policy interventions. Using data from 75,261 villages based on Indonesia’s Village Development Index (IDM), the research achieved exceptional results, with the Decision Tree model delivering a classification accuracy of 99.5%. Feature importance analysis identified the Economic Resilience Index (IKE) as the most influential factor, followed by the Ecological Resilience Index (IKL) and the Social Resilience Index (IKS), underscoring the critical role of these dimensions in rural development. These findings provide actionable insights for enhancing rural policies in Indonesia and contribute to a broader understanding of how machine learning can support sustainable development in similar contexts globally. By integrating advanced analytics with development strategies, this study offers a replicable framework to address rural-urban disparities and promote equitable, sustainable growth.

Keywords: Machine Learning; Artificial Intelligence; Underdeveloped Villages; Rural Development.

ARTICLE INFO
Received: May 14, 2024
Received in revised form: August 19, 2024
Accepted: December 25, 2024

doi: 10.46456/jisdep.v5i3.611

THE JOURNAL OF INDONESIA SUSTAINABLE DEVELOPMENT PLANNING
Published by Centre for Planners’ Development, Education, and Training (Pusbindiklatren), Ministry of National Development Planning/National Development Planning Agency (Bappenas), Republic of Indonesia

Address: Jalan Proklamasi 70, Central Jakarta, Indonesia 10320
Phone: +62 21 31928280/31928285
Fax: +62 21 31928281
E-mail: journal.pusbindiklatren@bappenas.go.id

Supported by Indonesian Development Planners Association (PPPI)

Please cite this article in APA Style as:
Al Azies, H. (2024). AI-Based Models for Identifying Underdeveloped Villages in Indonesia's Rural Development.*The Journal of Indonesia Sustainable Development Planning, Vol 5(3),* 192-202. https://doi.org/10.46456/jisdep.v5i3.611

1. Introduction

Rural development in Indonesia is crucial for advancing the nation's agenda toward achieving the Sustainable Development Goals (SDGs) (Castro-Arce & Vanclay, 2020). While urban areas have seen significant progress, rural regions continue to face enduring challenges such as inadequate infrastructure, including poorly maintained roads, unreliable electricity, and limited access to clean water (Otok et al., 2022). Additionally, access to essential services like education and healthcare remains restricted, exacerbating socioeconomic disparities between urban and rural populations (Parolin & Lee, 2021; Pujolar et al., 2022). Economic opportunities in these areas are often constrained by limited access to markets and financial services, hindering local livelihoods and economic growth (Aberese-Ako et al., 2022; Al Azies & Herowati, 2023; Cattaneo et al., 2022). These persistent issues highlight the substantial gap in development between urban and rural regions, underscoring the need for innovative approaches to bridge these disparities effectively (Bhatia et al., 2022).

To address these challenges, leveraging technology—particularly AI-based machine learning models—emerges as a promising solution for enhancing rural development outcomes(Ayoub Shaikh et al., 2022; Gikunda, 2024). By utilizing AI and machine learning, stakeholders can optimize resource allocation, improve decision-making processes, and identify targeted interventions tailored to local needs (Pakpahan, 2021). These technologies can analyze large datasets to uncover patterns and trends that inform development strategies, ensuring that interventions are relevant and impactful (Khair et al., 2020). Furthermore, engaging local communities in the implementation of these technologies fosters a sense of ownership and accountability, ultimately leading to more sustainable and inclusive growth (Lauwo et al., 2022). As Indonesia seeks to transform its rural landscapes, the integration of advanced technologies into development initiatives can play a pivotal role in driving equitable progress across the nation(Riggs et al., 2021; Wibowo, 2023).

In recent years, significant advancements have been made in utilizing machine learning methodologies to classify underdeveloped regions in Indonesia. For instance, studies by Palisoa et al. (2023) and Fadila Palisoa et al. (2020) have effectively demonstrated the capabilities of Support Vector Machine (SVM) and Radial Basis Function Neural Network (RBFNN) in identifying and analyzing the nuances of these areas. Research on the classification of underdeveloped regions in Indonesia has employed various machine learning methods to achieve high accuracy in identifying these areas. Palisoa et al. (2023) investigated the application of Support Vector Machine (SVM) for classifying underdeveloped regencies in Maluku Province, finding that SVM with a linear kernel function and parameter C=1 could correctly classify 76.13% of cases (Palisoa et al., 2023). Maulana and Irhamah (2018) also used SVM and Entropy-Based Fuzzy Support Vector Machine (EFSVM) for classifying regencies in East Java Province. Their results indicated that EFSVM performed better than SVM, especially with imbalanced data (Maulana & Irhamah, 2018). Fadila Palisoa et al (2020) applied Radial Basis Function Neural Network (RBFNN) for classifying underdeveloped regions in Indonesia, achieving a high performance with an accuracy of 93.48% (Fadila Palisoa et al., 2020). Sari et al. (2020) classified underdeveloped regencies in Eastern Indonesia using SVM with a linear kernel, reaching an accuracy of 87.23% (Sari et al., 2020). Putri et al (2023) proposed a novel approach for poverty mapping in East Java, Indonesia, using multi-source satellite imagery and point of interest data to create a more detailed poverty map at 1.5 km spatial resolution. Two scenarios were evaluated: one with zonal statistics from multi-source satellite imagery and geospatial data, and another using Resnet-34 transfer learning with daytime multiband and nighttime light intensity imagery. The CNN-1D model performed best in the first scenario, achieving an RMSE of 1.95 and an adjusted R^2 of 0.84 at the district level (Putri et al., 2023).

Hasugian et al. (2020) used the K-Means algorithm for classifying village status to support the Ministry of Village, Development of Disadvantaged Regions, and Transmigration programs. They successfully clustered 303 villages in Padang Regency based on their status (Hasugian et al., 2020). Utami and Wijayanto (2020) applied a Bayesian Network approach with K-Means Discretization for classifying the Village Development Index at the regency/city level, achieving an accuracy of 90.69% (Utami & Wijayanto, 2020). Otok et al. (2022) utilized an unsupervised learning approach for clustering underdeveloped infrastructure regions in Java, finding that the CLARA method provided the best results with high accuracy (Otok et al., 2022). Al Azies and Anuraga (2021) compared the SVM and K-Nearest

Neighbor (KNN) methods in classifying underdeveloped regions in Indonesia, discovering that both methods had the same precision value of 92.2% (Al Azies & Anuraga, 2021).

While these studies highlight the promise of machine learning in addressing rural underdevelopment, they also reveal notable gaps. Many approaches focus on a single algorithm or dataset, which limits their ability to capture the complex socio-economic dynamics of underdeveloped villages. Furthermore, the integration of multi-source data, including geospatial and socio-economic datasets, needs to be more utilized, constraining the potential for comprehensive analysis. Another critical gap is the need for more emphasis on translating these machine-learning insights into actionable policy recommendations, which are crucial for real-world impact. This research aims to address these limitations by employing a comprehensive framework that integrates multiple supervised learning algorithms, such as K-Nearest Neighbors (KNN), SVM, Logistic Regression, Decision Trees, and Naive Bayes. By analyzing diverse data sources and focusing on key socio-economic indicators, this study seeks to uncover nuanced patterns and causal relationships that traditional methods often overlook. Moreover, it emphasizes generating actionable insights to inform policy decisions, ensuring alignment with the SDGs. The novelty of this research lies in its integration of advanced machine learning techniques with diverse datasets to provide a granular analysis of underdevelopment at the village level. Unlike previous studies, this research not only focuses on achieving high accuracy in classification but also on delivering insights that can directly guide targeted and sustainable policy interventions. By incorporating community engagement in the implementation of these technologies, the study ensures inclusivity and long-term impact. Ultimately, this research aspires to set a benchmark for using AI in rural development, contributing to the reduction of rural-urban disparities and advancing Indonesia's sustainable development agenda.

2. Methods

This section outlines the comparative experimental methodology employed to analyze ensemble models in the classification of underdeveloped villages in Indonesia. The research specifically focuses on evaluating the performance of several supervised learning classification algorithms, including Decision Trees, Random Forests, Support Vector Machines (SVM), K-nearest neighbours (KNN), and Naive Bayes. Additionally, the effectiveness of various resampling techniques—such as oversampling and undersampling—on the dataset is examined to address class imbalance issues that may impact model performance. By implementing these methodologies, the study aims to identify the most suitable classification approach for accurately categorizing villages based on socio-economic indicators.

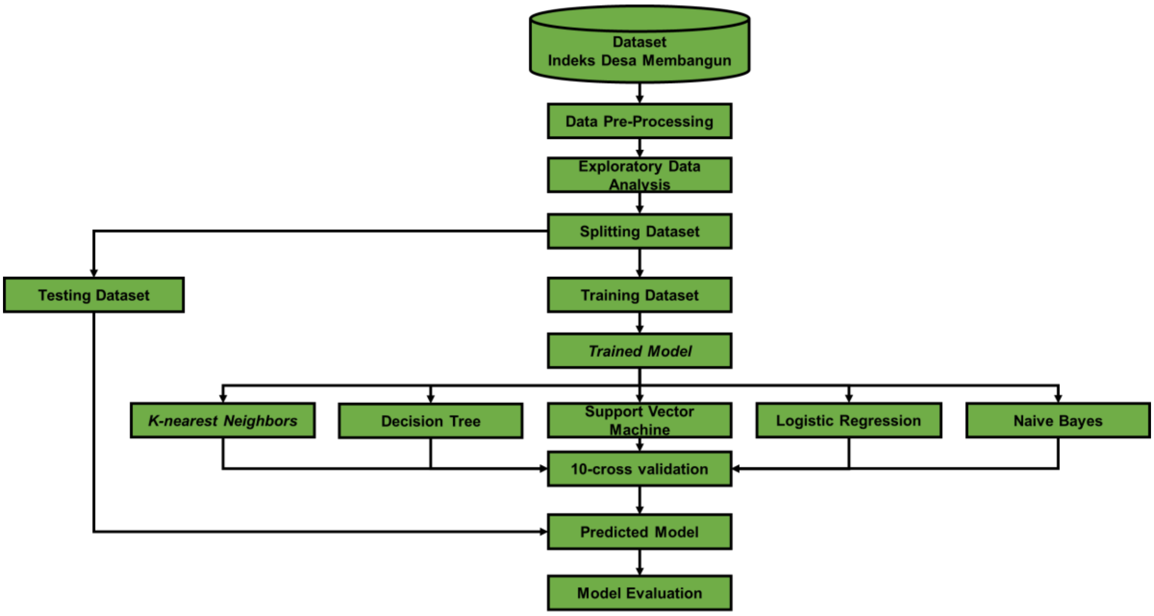

Figure 1. Research Framework for Comparative Experimental Approach to Classifying Underdeveloped Villages in Indonesia
Source: by Author

The research framework is illustrated in Figure 1, which depicts the stages of data preprocessing, model training, evaluation, and comparison. Data preprocessing involves cleaning the dataset, handling missing values, and applying feature selection techniques to ensure that only relevant variables are included in the analysis. Once the models are trained using the processed data, their performance is evaluated through metrics such as accuracy, precision, recall, F1-score, and confusion matrices. This comprehensive approach not only highlights the strengths and weaknesses of each algorithm but also provides insights into how different resampling methods can enhance classification outcomes. Ultimately, the goal is to derive actionable recommendations for policymakers and stakeholders engaged in rural development initiatives in Indonesia.

2.1 Data Sources and Research Variables

The data used in this research is sourced from the Ministry of Villages, Development of Disadvantaged Regions & Transmigration, specifically from Regulation 2016 Number 2 concerning the Village Development Index (IDM). The dataset comprises a comprehensive collection of information on 75,261 villages across Indonesia, reflecting their development status based on the IDM assessments conducted in 2023. The primary focus of this research is to classify these villages according to the Village Development Index, which serves as the target variable for analysis. Under the IDM framework, villages are categorized into five distinct classes: 'Advanced' (1), 'Independent' (2), 'Developing' (3), 'Underdeveloped' (4), and 'Very Underdeveloped' (5). This classification system offers a holistic appraisal of the development levels within each village, enabling researchers and policymakers to identify regions that require targeted interventions and particular emphasis on enhancement. By utilizing this structured classification, the study aims to contribute valuable insights into the current state of rural development in Indonesia and inform strategies for improving the socio-economic conditions in underdeveloped and very underdeveloped villages. This research employs the three components that comprise the Village Development Index (IDM) from the Ministry of Villages, Development of Disadvantaged Regions & Transmigration, which in these concerns where:

a. **Social Resilience Index (IKS)**
   Measures the social resilience of a village, covering aspects such as education, health, and social welfare.
b. **Economic Resilience Index (IKE)**
   Evaluates the economic resilience of a village, involving economic parameters such as per capita income, employment opportunities, and access to economic resources.
c. **Ecological/Environmental Resilience Index (IKL)**
   Reflects the environmental resilience of a village, including factors such as natural resource management, environmental sustainability, and environmental impact.

2.2 Data Preprocessing

The analysis stages begin with exploratory data analysis (EDA) to understand the distribution and relationships of the variables (Indrakumari et al., 2020). This consists of calculating descriptive statistics and creating images in order to look for patterns, trends, and outliers in the data. Exploratory Data Analysis (EDA) is used in analyzing the data structure and this is done before the modeling stage. Then subsequently outlier detection and handling are performed considering that the extreme values can introduce bias in the model. Outliers can bias statistical measures and affect models' performance hence the reason why they need to be effectively handled. Z-scores, IQR, or robust scaling can be applied to find and remedy outliers. Z-score standardization has always been there before other standardization types such as range or min-max standardization where an effort is made to adjust the domains of the variables in an IT. Purely means transforming or normalizing the above information would give Z = 0, S = 1 therefore the use of such algorithms like SVM and KNN which are very susceptible to data size is highly sensitive. Finally, the main dataset is divided duplicated into a composite training data set and a held-out test data set for assessing performance of the model built on the 'training' data using 'unseen' data (Bates et al., 2023).

2.3. Development of Classification Models

AI and ML are two significant elements that provide opportunities to solve and comprehend complex problems, such as the development disparity of poor villages (Figure 2A). AI is the capacity of machines or computer programs to undertake activities that generally demand human intelligence (Azies & Rositawati, 2021; Santoso & Kom, 2023). In this line of thought, AI has various components, including ML (Figure 2A). The critical difference is that M L sets an objective and develops models and algorithms that let the machine interpret the data independently without the need to be programmed. Concerning underdeveloped villages, ML can assist in developing predictive models that classify and interpret the figures of underdevelopment. Moreover, this will also help to determine some of the significant factors that hinder the village's development and enable better decisions to be made (Santoso & Kom, 2023). Including AI and ML in assessing underdeveloped villages brings a new understanding of the data. According to many definitions, supervised learning is a paradigm of Machine Learning wherein models or algorithms are developed by training on pre-classified data (Silalahi et al., 2023). Supervised Learning is an approach in Machine Learning where models or algorithms are trained using labeled data (Silalahi et al., 2023). In this context, "labels" refer to the desired or known outputs. The main goal of Supervised Learning is to learn the relationships or patterns between inputs and outputs so that the model can make accurate predictions or classifications on new, unseen data (Brnabic & Hess, 2021). There are various types of Supervised Learning algorithms (Figure 2B).

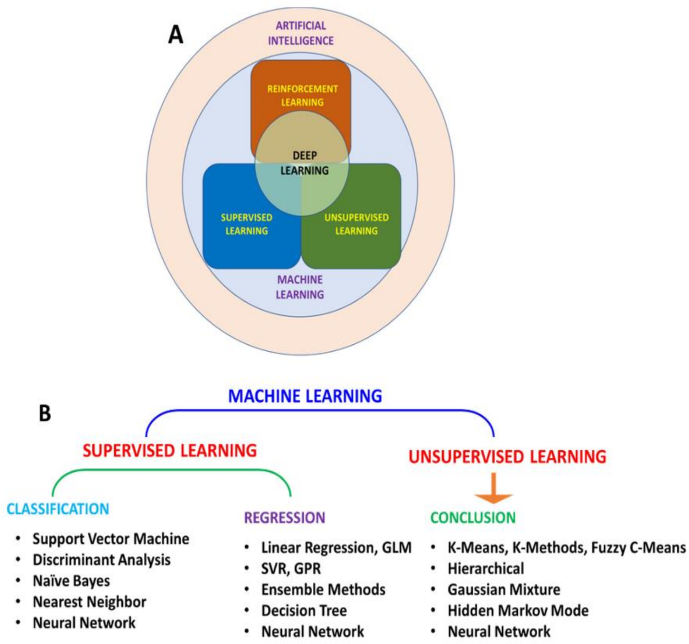

Figure 2. Representation of the schema relationship between AI, ML, and DL (A); classification of ML algorithms (B) (Manickam et al., 2022)

2.4. Evaluation of Classification Models

Classification models are defined, and rules that will underlie the predictions are produced as they are used. These rules must be checked and verified to assess the level of prediction correctness. The classification rules evaluation and validation are done with the help of a Confusion matrix (Deng et al., 2016),, which helps summarize the classification model's performance. In this case, metrics obtained from the confusion matrix include accuracy, which refers to how many instances were classified correctly concerning the total number of classified instances. Another metric is precision, whereby the proportion of actual positive instances out of the ones predicted to be positive is evaluated. Recall, called sensitivity, defines the ratio of accurate positive and positive results. The F1 score normatively covers the gap between the precision and recall scores to get an intermediate assessment.

3. Results and Discussions

3.1. Overview of Underdeveloped Areas in Indonesia

In this subsection, the overall information concerning the scenario of backward villages in Indonesia comes through descriptive statistical analysis. The explanation and graphical representations of the Social Resilience Index, Economic Resilience Index, and Ecological/Environmental Resilience Index are carried out. These descriptive statistics are then utilized by standard learning algorithms, as these characteristics of underdeveloped villages are understood before making them divided using supervised learning algorithms.

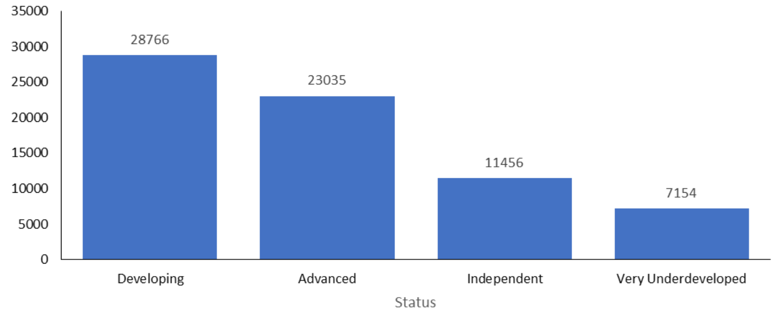

Figure 3. Distribution of Villages by Development Status in Indonesia

In Indonesia, as shown in Figure 3, the data shows that developing villages constitute a significant portion of the total, with 28,766 of the 75,261 villages studied classified as developing villages. Additionally, there are 23,035 advanced villages, 11,456 independent villages, 4,850 very underdeveloped villages, and 7,154 underdeveloped villages. While the categories of creating and advanced villages dominate the overall distribution, it is crucial to note that the very underdeveloped and underdeveloped categories still represent a substantial number of villages. The Social Resilience Index (IKS) measures various factors, including the level of education attained by individuals, the availability of healthcare services, community participation in social activities, and the overall safety and comfort experienced within the village. This index provides valuable context for understanding the social dynamics in these areas and serves as a foundational metric for evaluating the resilience and development status of villages across Indonesia. By analyzing the IKS, IKE, and IKL, researchers can gain insights into the underlying challenges faced by backward villages and the factors contributing to their development status. These descriptive statistics will guide the implementation of machine learning algorithms, helping to classify villages based on their resilience indices and other socio-economic indicators. This approach aims to uncover patterns and relationships that can inform targeted interventions and policies to promote rural development in Indonesia.

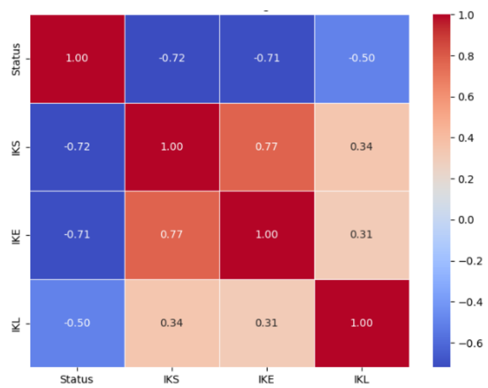

Figure 4. Correlation Matrix Between Village Status and Social, Economic, and Ecological Resilience Indices Source

The correlation analysis (Figure 4) reveals several key insights about the relationship between village status and the indices of resilience:

- **Village Status and IKS, IKE, and IKL**: The correlation between village status and IKS is -0.720250, indicating that villages with better status tend to have higher IKS values. This means that more advanced villages have better social resilience. Similarly, the correlation between village status and IKE is -0.711490, suggesting that villages with better status also tend to have higher IKE values, indicating better economic resilience in more advanced villages. The correlation between village status and IKL is -0.504073, which shows that villages with better status tend to have better ecological resilience, although this relationship is not as strong as with IKS and IKE.
- **Social Resilience Index (IKS) with IKE and IKL**: The correlation between IKS and IKE is 0.773192, showing a strong positive relationship. Villages with good social resilience tend to also have good economic resilience. The correlation between IKS and IKL is 0.343444, indicating a positive but weaker relationship compared to the correlation between IKS and IKE.
- **Economic Resilience Index (IKE) with IKL**: The correlation between IKE and IKL is 0.309940, showing a positive but weak relationship. Villages with good economic resilience tend to have good ecological resilience, although this relationship is not as strong as the relationship between social and economic resilience.

These correlations highlight the interconnectedness of social, economic, and ecological resilience in determining the overall status of villages, with social and economic resilience being more strongly linked than ecological resilience.

3.2. Classification Models for Underdeveloped Areas Using Machine Learning

This section discusses the performance of various machine learning models used to classify underdeveloped villages in Indonesia. The models evaluated include K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression, Decision Tree, and Naive Bayes. The performance metrics considered are accuracy, precision, recall, and F1-score. The results are summarized in the Table 1 below.

Table 1: Performance Metrics of Classification Models for Underdeveloped Villages.

Algorithm	Accuracy	Precision	Recall	F1-score
KNN	0.994752	0.994753	0.994752	0.994746
SVM	0.991364	0.991426	0.991364	0.991355
Logistic Regression	0.990500	0.990580	0.990500	0.990479
Decision Tree	0.995483	0.995482	0.995483	0.995479
Naïve Bayes	0.878031	0.881352	0.878031	0.877753

Source: Processed data by Python (2024)

In the evaluation of models for classifying underdeveloped villages, the Decision Tree model emerges as the top performer with an impressive accuracy of 0.995483. This model not only achieves high accuracy but also demonstrates exceptional precision, recall, and F1-score metrics, highlighting its robustness in accurately identifying underdeveloped areas. Following closely, the K-Nearest Neighbors (KNN) model achieves an accuracy of 0.994752, showing strong performance across all metrics. This reinforces its effectiveness as another reliable model for this classification task. The Support Vector Machine (SVM) model, with an accuracy of 0.991364, also performs well, maintaining strong precision, recall, and F1-score metrics, indicating its suitability for the task. Logistic Regression similarly shows robust performance with an accuracy of 0.990500, aligning well with the other top-performing models.

In contrast, the Naive Bayes model exhibits a notably lower accuracy of 0.878031. While its precision, recall, and F1-score metrics are consistent with its accuracy, these results underscore its relative limitations compared to the Decision Tree, KNN, SVM, and Logistic Regression models. Overall, the results highlight the effectiveness of machine learning models, particularly the Decision Tree model, in accurately classifying underdeveloped villages. Its high accuracy and comprehensive performance metrics make it a valuable tool for identifying and addressing underdevelopment challenges in Indonesia, supporting informed policy decisions and targeted interventions in rural areas.

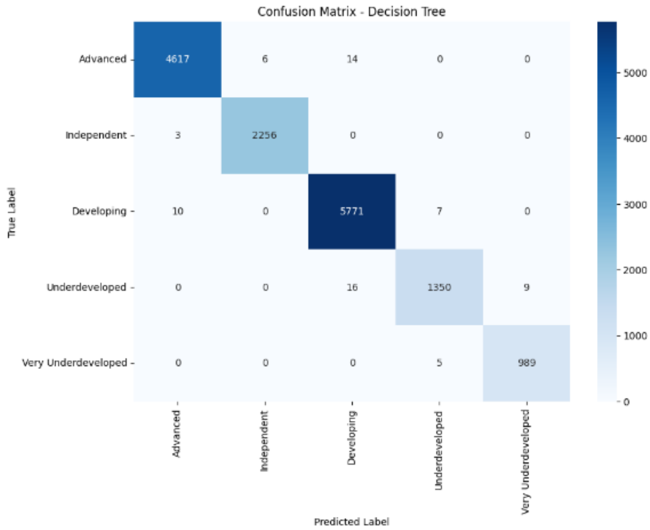

Figure 5. Confusion Matrix Analysis for Decision Tree Model in Classifying Underdeveloped Villages

The selected Decision Tree model exhibits an outstanding accuracy of 99.55%, highlighting its exceptional ability to classify underdeveloped villages accurately. This high accuracy reflects the model's effectiveness in distinguishing between different development statuses. The accompanying confusion matrix (Figure 5) provides further insights into the model's performance. It indicates that a substantial majority of actual underdeveloped villages are accurately identified, as evidenced by a high True Positive rate. Moreover, the model demonstrates a deficient number of False Negatives, which are villages that are genuinely underdeveloped but incorrectly predicted as not underdeveloped. This minimal occurrence signifies that the model is highly reliable in capturing the villages that require attention and support.

In addition to high accuracy, the Decision Tree model maintains a well-balanced performance across key metrics, including precision, recall, and F1-score. These balanced values reinforce the model's robustness in accurately identifying and classifying underdeveloped villages across Indonesia. The strong performance of the Decision Tree model positions it as a valuable tool for stakeholders aiming to effectively identify and address developmental challenges in these regions. Its reliable predictions can facilitate targeted interventions, enabling policymakers and development agencies to allocate resources and design programs that address the specific needs of underdeveloped villages.

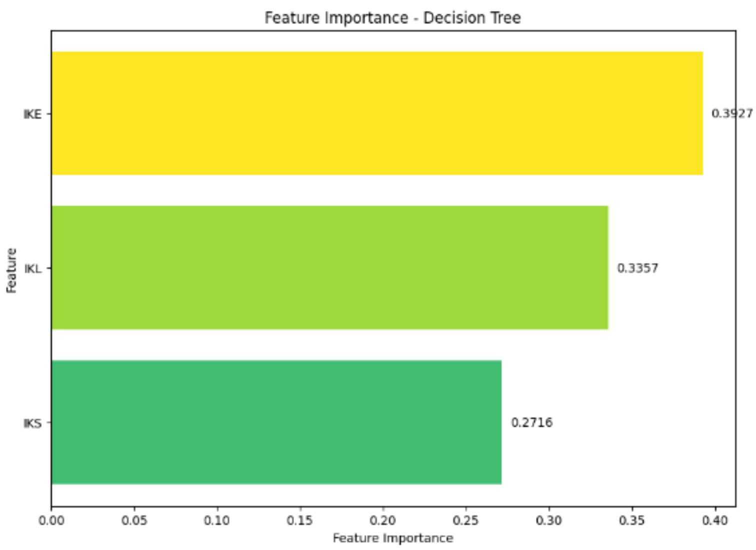

Figure 6. Feature Importance Analysis of Decision Tree Model for Classification of Underdeveloped Villages

After discussing the Confusion Matrix results that provide an overview of the model's performance in After discussing the Confusion Matrix results that provide an overview of the model's performance in predicting underdeveloped villages, we further explore the Feature Importance of the Decision Tree Model, as presented in Figure 6. Feature Importance indicates the extent to which each feature or variable contributes to the model's ability to classify. In this context, the Feature Importance from the Decision Tree Model illustrates how significant the Social Resilience Index (IKS), Economic Resilience Index (IKE), and Ecological/Environmental Resilience Index (IKL) are in predicting whether a village falls into the categories of "Underdeveloped" or "Very Underdeveloped." Based on Figure 6, it is observed that the IKE feature has the highest Feature Importance, followed by IKL and IKS. This suggests that in the classification of underdeveloped villages, economic resilience (IKE) has the most substantial influence on predictions, followed by ecological/environmental resilience (IKL) and social resilience (IKS). Therefore, focusing on these aspects in the analysis helps in understanding and identifying key factors influencing the underdeveloped status of villages in Indonesia

Conclusion

This study explored the application of machine learning techniques to predict and classify underdeveloped villages in Indonesia, focusing on the Social Resilience Index (IKS), Economic Resilience Index (IKE), and Ecological/Environmental Resilience Index (IKL). By employing Decision trees and other supervised learning models, the research achieved high accuracy in identifying underdeveloped areas, offering valuable insights into the factors influencing rural development. Economic resilience (IKE) emerged as the most significant predictor, followed by ecological/environmental resilience (IKL) and social resilience (IKS). These findings underscore the critical role of economic and environmental factors in shaping rural development and provide a data-driven foundation for formulating targeted interventions to address underdevelopment. The results have direct practical relevance for policymakers aiming to align rural development initiatives with the Sustainable Development Goals (SDGs). By identifying key resilience factors, this study provides actionable insights that can guide evidence-based strategies to reduce disparities, improve livelihoods, and foster sustainable progress in rural communities.

Limitations

Despite its contributions, this study has several limitations. The predictive accuracy of the models may vary across different geographical regions and temporal contexts due to the dynamic nature of socioeconomic and environmental factors. The reliance on secondary data from national surveys may introduce biases or inaccuracies inherent in such datasets. Additionally, the study’s scope was confined to specific resilience indices, potentially overlooking other influential factors, such as cultural dynamics or local governance structures. Future research could address these limitations by incorporating more diverse and comprehensive datasets, including qualitative data on cultural and governance dynamics. Enhancing model robustness through ensemble techniques or spatial-temporal analyses could further improve predictive accuracy and applicability. By broadening the scope and refining methodologies, subsequent studies can build on these findings to create more holistic and effective strategies for rural development in Indonesia. This research highlights the potential of machine learning as a transformative tool in the pursuit of SDG-aligned policymaking. By leveraging data-driven insights, stakeholders can design targeted, impactful interventions that empower rural communities and ensure no one is left behind in Indonesia’s journey toward sustainable development.

References

Aberese-Ako, M., Immurana, M., Dalaba, M. A., Anumu, F. E. Y., Ofosu, A., & Gyapong, M. (2022). The socio-economic and health effects of COVID-19 among rural and urban-slum dwellers in Ghana: A mixed methods approach. *PLOS ONE*, *17*(7), e0271551. https://doi.org/10.1371/JOURNAL.PONE.0271551

Al Azies, H., & Anuraga, G. (2021). Classification of Underdeveloped Areas in Indonesia Using the SVM and k-NN Algorithms. *Jurnal ILMU DASAR*, *22*(1), 31–38. https://doi.org/10.19184/JID.V22I1.16928

Al Azies, H., & Herowati, W. (2023). Unravelling Income Inequality in Indonesia. *Jurnal Riset Ilmu Ekonomi*, *3*(2), 89–100. https://doi.org/10.23969/JRIE.V3I2.63

Ayoub Shaikh, T., Rasool, T., & Rasheed Lone, F. (2022). Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. *Computers and Electronics in Agriculture*, *198*, 107119. https://doi.org/10.1016/J.COMPAG.2022.107119

Azies, H. Al, & Rositawati, A. F. D. (2021). Mapping of the Reading Literacy Activity Index in East Java Province, Indonesia: an Unsupervised Learning Approach. *Proceedings of The International Conference on Data Science and Official Statistics*, *2021*(1), 211–223. https://doi.org/10.34123/ICDSOS.V2021I1.128

Bates, S., Hastie, T., & Tibshirani, R. (2023). Cross-Validation: What Does It Estimate and How Well Does It Do It? *Https://Doi.Org/10.1080/01621459.2023.2197686*. https://doi.org/10.1080/01621459.2023.2197686

Bhatia, S., Landier, W., Paskett, E. D., Peters, K. B., Merrill, J. K., Phillips, J., & Osarogiagbon, R. U. (2022). Rural–Urban Disparities in Cancer Outcomes: Opportunities for Future Research. *JNCI: Journal of the National Cancer Institute*, *114*(7), 940–952. https://doi.org/10.1093/JNCI/DJAC030

Brnabic, A., & Hess, L. M. (2021). Systematic literature review of machine learning methods used in the analysis of real-world data for patient-provider decision making. *BMC Medical Informatics and Decision Making*, *21*(1), 1–19. https://doi.org/10.1186/S12911-021-01403-2/FIGURES/2

Castro-Arce, K., & Vanclay, F. (2020). Transformative social innovation for sustainable rural development: An analytical framework to assist community-based initiatives. *Journal of Rural Studies*, *74*, 45–54. https://doi.org/10.1016/J.JRURSTUD.2019.11.010

Cattaneo, A., Adukia, A., Brown, D. L., Christiaensen, L., Evans, D. K., Haakenstad, A., McMenomy, T., Partridge, M., Vaz, S., & Weiss, D. J. (2022). Economic and social development along the urban–rural continuum: New opportunities to inform policy. *World Development*, *157*, 105941. https://doi.org/10.1016/J.WORLDDEV.2022.105941

Deng, X., Liu, Q., Deng, Y., & Mahadevan, S. (2016). An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Information Sciences*, *340–341*, 250–261. https://doi.org/10.1016/J.INS.2016.01.033

Fadila Palisoa, N., Sinay, L. J., Matdoan, M. Y., & Bakarbessy, L. (2020). Penerapan Radial Basis Function Neural Network dalam Pengklasifikasian Daerah Tertinggal di Indonesia. *Jurnal Aplikasi Statistika & Komputasi Statistik*, *12*(1), 37–54. https://doi.org/10.34123/JURNALASKS.V12I1.250

Gikunda, K. (2024). *Harnessing Artificial Intelligence for Sustainable Agricultural Development in Africa: Opportunities, Challenges, and Impact*. https://arxiv.org/abs/2401.06171v1

Hasugian, P. M., Hutahaean, H. D., Sinaga, B., Sriadhi, & Silaban, S. (2020). Villages Status Classification Analysis Involving K-Means Algorithm To Support Kementerian Desa Pembangunan Daerah Tertinggal dan Transmigrasi Work Programs. *Journal of Physics: Conference Series*, *1641*(1), 012058. https://doi.org/10.1088/1742-6596/1641/1/012058

Indrakumari, R., Poongodi, T., & Jena, S. R. (2020). Heart Disease Prediction using Exploratory Data Analysis. *Procedia Computer Science*, *173*, 130–139. https://doi.org/10.1016/J.PROCS.2020.06.017

Khair, N. K. M., Lee, K. E., & Mokhtar, M. (2020). Sustainable City and Community Empowerment through the Implementation of Community-Based Monitoring: A Conceptual Approach. *Sustainability 2020, Vol. 12, Page 9583*, *12*(22), 9583. https://doi.org/10.3390/SU12229583

Lauwo, S. G., Azure, J. D. C., & Hopper, T. (2022). Accountability and governance in implementing the Sustainable Development Goals in a developing country context: evidence from Tanzania. *Accounting, Auditing and Accountability Journal*, *35*(6), 1431–1461. https://doi.org/10.1108/AAAJ-10-2019-4220/FULL/PDF

Manickam, P., Mariappan, S. A., Murugesan, S. M., Hansda, S., Kaushik, A., Shinde, R., & Thipperudraswamy, S. P. (2022). Artificial Intelligence (AI) and Internet of Medical Things (IoMT)

Assisted Biomedical Systems for Intelligent Healthcare. *Biosensors 2022, Vol. 12, Page 562*, *12*(8), 562. https://doi.org/10.3390/BIOS12080562

Maulana, J. P., & Irhamah, I. (2018). Klasifikasi Kabupaten di Provinsi Jawa Timur Berdasarkan Indikator Daerah Tertinggal dengan metode Support Vector Machine (SVM) dan Entropy Based Fuzzy Support Vector Machine (EFSVM). *Inferensi*, *1*(1), 9–15. https://doi.org/10.12962/J27213862.V1I1.6715

Otok, B. W., Suharsono, A., Purhadi, Standsyah, R. E., & Azies, H. Al. (2022). Partitional Clustering of Underdeveloped Area Infrastructure with Unsupervised Learning Approach: A Case Study in the Island of Java, Indonesia. *Journal of Regional and City Planning*, *33*(2), 177–196. https://doi.org/10.5614/JPWK.2022.33.2.3

Pakpahan, R. (2021). Analisa Pengaruh Implementasi Artificial Intelligence Dalam Kehidupan Manusia. *JISICOM (Journal of Information System, Informatics and Computing)*, *5*(2), 506–513. https://doi.org/10.52362/JISICOM.V5I2.616

Palisoa, N. F., Sinay, L. J., Matdoan, M. Y., Yudistira, Y., & Bakarbessy, L. (2023). PENERAPAN SUPPORT VECTOR MACHINE (SVM) UNTUK KLASIFIKASI KABUPATEN TERTINGGAL DI PROVINSI MALUKU. *PARAMETER: Jurnal Matematika, Statistika Dan Terapannya*, *2*(02), 79–86. https://doi.org/10.30598/PARAMETERV2I02PP79-86

Parolin, Z., & Lee, E. K. (2021). Large socio-economic, geographic and demographic disparities exist in exposure to school closures. *Nature Human Behaviour 2021 5:4*, *5*(4), 522–528. https://doi.org/10.1038/s41562-021-01087-8

Pujolar, G., Oliver-Anglès, A., Vargas, I., & Vázquez, M. L. (2022). Changes in Access to Health Services during the COVID-19 Pandemic: A Scoping Review. *International Journal of Environmental Research and Public Health 2022, Vol. 19, Page 1749*, *19*(3), 1749. https://doi.org/10.3390/IJERPH19031749

Putri, S. R., Wijayanto, A. W., & Pramana, S. (2023). Multi-source satellite imagery and point of interest data for poverty mapping in East Java, Indonesia: Machine learning and deep learning approaches. *Remote Sensing Applications: Society and Environment*, *29*, 100889. https://doi.org/10.1016/J.RSASE.2022.100889

Riggs, R. A., Achdiawan, R., Adiwinata, A., Boedhihartono, A. K., Kastanya, A., Langston, J. D., Priyadi, H., Ruiz-Pérez, M., Sayer, J., & Tjiu, A. (2021). Governing the landscape: potential and challenges of integrated approaches to landscape sustainability in Indonesia. *Landscape Ecology*, *36*(8), 2409–2426. https://doi.org/10.1007/S10980-021-01255-1/METRICS

Santoso, J. T., & Kom, M. (2023). KECERDASAN BUATAN (Artificial Intelligence). *Penerbit Yayasan Prima Agus Teknik*, 1–227. https://penerbit.stekom.ac.id/index.php/yayasanpat/article/view/437

Sari, E. A., Saragih, M. T. Br., Shariati, I. A., Sofyan, S., Baihaqi, R. Al, & Nooraeni, R. (2020). Klasifikasi Kabupaten Tertinggal di Kawasan Timur Indonesia dengan Support Vector Machine. *JIKO (Jurnal Informatika Dan Komputer)*, *3*(3), 188–195. https://doi.org/10.33387/JIKO.V3I3.2364

Silalahi, A. P., Simanullang, H. G., & Hutapea, M. I. (2023). SUPERVISED LEARNING METODE K-NEAREST NEIGHBOR UNTUK PREDIKSI DIABETES PADA WANITA. *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, *7*(1), 144–149. https://doi.org/10.46880/JMIKA.VOL7NO1.PP144-149

Utami, N. A., & Wijayanto, A. W. (2020). Classification of Village Development Index at Regency/Municipality Level Using Bayesian Network Approach with K-Means Discretization. *Jurnal Aplikasi Statistika & Komputasi Statistik*, *12*(3), 95–106. https://doi.org/10.34123/JURNALASKS.V14I1.390

Wibowo, A. (2023). Global Economic Challenges for Indonesia: Equitable Development for Sustainable Prosperity. *Prosiding Seminar Teknologi, Akuntansi, Bisnis, Ekonomi, Dan Komunitas*, *3*, 23–40. https://doi.org/10.35912/STABEK.V3I.175