

Clustering Provincial Democratic Performance in Indonesia Using Machine Learning

Journal of Eurasian Studies
2026, Vol. 17(2) 481–491

Mochammad Rozikin¹, Muhamad Heri Romadhon¹, Andy Fefta Wijaya², Fadillah Amin², Alfi Haris Wanto², Andjar Prasetyo³, Mardiyono⁴, Hartiningsih⁵, Untung Muhdiarta⁶, Heri Wahyudianto⁷

Abstract

This study examines the regional variation of the Indonesian Democracy Index (IDI) across 34 provinces from 2021 to 2024 in the context of global democratic backsliding. Understanding spatial disparities in democracy is crucial for promoting inclusive and evidence-based governance. Using the Random Forest Clustering (RFC) method, provinces were grouped based on indicators of civil liberties, political institutions, and community participation. Model selection was optimized by the Bayesian Information Criterion (BIC), with performance evaluated through R^2 , silhouette score, and the Calinski–Harabasz index. The optimal model produced four clusters ($R^2 = 0.806$; silhouette = 0.310), revealing significant spatial diversity in democratic quality. Cluster 1 represents provinces with moderate stability and high heterogeneity; Cluster 2 includes those with limited democracy and low participation; Cluster 3 demonstrates consistent and balanced democratic performance; and Cluster 4 contains provinces with cohesive and mature democratic institutions. These results highlight interregional inequalities in democracy linked to institutional capacity and civic engagement. At the national level, the findings guide more targeted policy interventions for democratic strengthening. At the global level, this research contributes to understanding democratic resilience in developing countries and demonstrates the analytical value of machine learning for comparative studies of governance and political systems

Keywords

Democracy, clusters, machine learning, Indonesian Democracy Index, global issues.

Introduction

Democracy is one of the main indicators of the quality of modern governance because it guarantees civil liberties, political accountability, and a strong power control mechanism between state institutions and citizens. However, in recent years the phenomenon of democratic backsliding, that is, the decline in the quality of democracy through the weakening of civil liberties, the rule of law, and political accountability, has emerged in many countries in different parts of the world, showing that democracy does not always move forward in a linear manner but can undergo regression (Hellmeier et al., 2021). International studies show a global trend in which the right to freedom of expression, the rule of law, and political competition have weakened in the last decade, so that the number of countries showing signs of democratic erosion continues to increase even though some countries maintain a relatively free electoral process (Little & Meng, 2024). This phenomenon is also positioned as part of a global challenge to contemporary democracy, where there is debate about the nature, size, and consequences of the declining trend in the

quality of democracy in the post-1990 era (Bartels et al., 2023). The literature states that backsliding can occur through elected government actions that gradually weaken power-balancing institutions such as parliament, independent media, and the judiciary to strengthen executive control (Hellmeier et al., 2021). In addition, the

¹Doctoral Program in Administrative Sciences, Faculty of Public Administration, Universitas Brawijaya, Malang, Indonesia

²Department of Public Administration at the Faculty of Administrative Sciences, Brawijaya University, Indonesia

³Regional Development Planning, Research and Innovation Agency of Magelang City, Central Java Province, Indonesia

⁴Center for Population Research, National Research and Innovation Agency, Jakarta, Indonesia

⁵Regional Research and Innovation Agency of South Kalimantan Province, Banjarbaru City, South Kalimantan Province, Indonesia

⁶Faculty of Social and Political Sciences, Cendrawasih University, Jayapura City, Papua Province, Indonesia

⁷Regional Development Planning, Research and Innovation Agency of Papua Province, Jayapura City, Papua Province, Indonesia



relationship between democratic performance and the provision of public services is an important study material in understanding public trust in the modern democratic system: government performance in meeting economic and social outcomes can strengthen democratic legitimacy or conversely accelerate public dissatisfaction if the democratic system fails to show results (Fukuyama et al., 2025). A multidimensional approach to democracy emphasizes that democracy is not only about electoral competition, but also inclusivity, protection of civil rights and the effectiveness of government (Hellmeier et al., 2021). Empirical evidence from global comparative studies shows that although not all indicators of democracy show a uniform downward trend, there have been substantial declines in several dimensions such as media freedom and judicial law in various countries, suggesting that the quality of democracy in aggregate is under significant pressure (Little & Meng, 2024). Cross-border analysis also shows the complexity of the phenomenon of democracy in the contemporary era: while some countries are experiencing stagnation or improvement in the quality of democracy, many others are showing signs of institutional weakness amid increasing political polarization and other internal challenges (Bartels et al., 2023). In the Indonesian context, this situation demands serious attention to the dynamics of democracy at the sub-national level, as similar challenges can arise in the form of weaknesses in political pluralism, social fragmentation and pressures on civil liberties if not anticipated through inclusive, data-driven and responsive democratic policies to socio-political change. Thus, a deeper understanding of democratic patterns in each province will provide an empirical basis to support strong and sustainable democratic policies, while also considering the complexity of diverse global democratic challenges.

Although Indonesia's democracy has evolved since the post-reform era, empirical evidence suggests that the quality of democracy at the subnational level tends to be uneven between regions, where provinces with strong political institutions often differ significantly from provinces that are still weak in public participation, civil liberties, and political accountability. In the contemporary democracy literature, researchers emphasize that the spatial variation of democracy within the country is an important phenomenon that cannot be ignored, as the democratization process often operates differently at different levels of government and territory (Gagnon & Fleuss, 2020). The study of the concept of subnational democracy also emphasizes that the quality of democracy, including the regularity of power change, competition, and local autonomy, can vary greatly between regions even within a country, demonstrating the importance of measuring political freedom and

decentralized democratic practices to understand local dynamics (Harbers et al., 2019). This variation is influenced by the distribution of political power, regional institutional structures, and the complex and multilevel governance relationship between central and regional governments that affect the way democracy is implemented in different regions (OECD, 2022). Comparative research also shows that the spatial differentiation of democracy is a global phenomenon, where the quality of democracy can differ significantly between urban and rural areas or between regions with different bureaucratic resources and capacities, which further impacts public participation and political control (Boese et al., 2022). In the context of contemporary democracy, global studies also show that multilevel democracy is backsliding, namely changes in the quality of democratic practices that occur not only at the national level, but also at the local and regional levels that have an impact on the overall cohesion of democracy (Little & Meng, 2024). Other studies affirm that subnational analysis provides important insights into how democratic institutions are changing and why some regions are developing more politically stable while others remain lagging behind, due to structural factors such as political autonomy, bureaucratic quality, and geographically different community participation (Gagnon & Fleuss, 2020). Other literature underscores that geographic variation in democracy is rooted not only in structural differences but also in local capacities to maintain civil liberties, control of corruption, and the effectiveness of governance (Pavel, 2025). Taking this approach into account, the study on the IDI for the period 2021–2024 aims to identify the spatial patterns of democracy, so that the mapping of the quality of democracy not only assesses national performance but also captures local variations that have strong implications for inclusive, responsive, and evidence-based democratic policies.

Democracy is one of the main indicators of the quality of modern governance, but in recent years the phenomenon of democratic backsliding, namely the decline in the quality of democracy due to the weakening of civil liberties, political accountability, and the decline in the functioning of democratic institutions, has become the focus of studies in various national and international contexts (Liu et al., 2021). Contemporary studies show that the decline in the quality of democracy does not only occur at the national level but also has important implications at the sub-national level when there are inequalities in political institutions, public participation, and civil liberties (Schneider, 2018). In this global context, Indonesia, as the largest democracy in Southeast Asia, faces similar challenges at the national and regional levels, where spatial variation in interprovincial democracy indicators is critical to understand in efforts to

strengthen more inclusive and evidence-based democratic policies. Although Indonesia's democracy has developed rapidly after the Reformation, its quality has not been evenly distributed between regions; Some provinces exhibit strong political institutions, high public participation, and relatively well-maintained civil liberties, while others still face significant limitations in these aspects. This kind of spatial variation is noted in the political geography literature that emphasizes the importance of regional and structural dynamics in determining the quality of local democracy (Kluser & Breunig, 2023). In addition, cross-country studies show that the strength of local institutions, administrative capacity, and political information infrastructure of each region significantly affect the effectiveness of democratic processes (Adam et al., 2022). The main challenge of this study is to identify how these spatial variations are formed and determine their grouping patterns in the context of the IDI for the 2021–2024 period, so as to support prescriptive and contextual policy recommendations. Most previous studies have assessed the quality of democracy in aggregate at the national level, without taking into account the spatial dimensions and inter-regional linkages within a country. The sub-national and regional democratic literature states that measurements that take into account geographic context and interregional interactions provide sharper insights into dynamics in democratic systems (Grubman, 2022). However, there is still an academic gap related to the use of advanced quantitative methods such as machine learning to uncover latent structures in democracy data that are multivariate and closely related to spatial factors. Traditional approaches such as descriptive statistics or linear regression often fail to uncover complex and non-linear clustering patterns in big data (Waggoner, 2020). Therefore, a more robust and objective analytical approach, such as Random Forest Clustering, which harnesses the power of ensemble learning to discover the structure and typology of regional democracy without rigid parametric assumptions, is needed. The application of clustering algorithms in contemporary social science research has been theoretically supported to identify patterns in complex and highly high-dimensional data (Waggoner, 2020), as well as to map variation in behavior or social structure at the individual and group levels (Benati et al., 2025). In addition, unsupervised learning methods such as clustering allow the grouping of observations based on multidimensional similarities without labels, which is suitable for extracting the typology of democratic regimes or the quality of local democracy based on available index data (Waggoner, 2020). Through this kind of approach, this study aims not only to identify provincial groups with similar democratic characteristics, but also to uncover the key attributes that distinguish these groups empirically so as to provide a

stronger basis for responsive and targeted democratic policymaking. This research contributes in two main domains: scientifically, introducing a data-driven approach to democracy analysis; Practically, it provides a more directed policy basis for the government in the equitable distribution of democratic quality. Globally, this research also enriches comparative studies of the variation and resilience of democracy in developing countries.

Method

This study uses an exploratory quantitative (Gosiewska & Biecek, 2020) approach with the RFC method to identify interprovincial grouping patterns based on the IDI. This unsupervised machine learning approach was chosen because it is able to recognize non-linear patterns and complex relationships between indicators without the need for dependent variables. RFCs are also more resistant to noise and multicollinearity than conventional clustering methods such as k-means or hierarchical clustering.

To improve methodological transparency and reproducibility, the RFC procedure is described in terms of data preparation, forest construction, proximity estimation, cluster formation, model evaluation, and visualization. These stages distinguish the generation of the similarity structure from the subsequent assignment of provinces to clusters.

The research was conducted in 34 provinces in Indonesia using secondary IDI data for 2021–2024 published by the Central Statistics Agency (BPS, 2025). This period was chosen because it represents an important phase post-pandemic and ahead of the 2024 elections, which strategically affects the dynamics of politics and democratic governance at the regional level. This study uses three main variables of IDI which are predictive (independent) and operationalized based on the official methodology of (BPS, 2025).

The province constitutes the unit of analysis, covering 34 provinces in Indonesia. For each province, the annual scores of the three IDI dimensions for 2021–2024 were averaged to obtain one representative score for each dimension over the study period. Consequently, the analytical dataset consisted of 34 provincial observations and three input variables. Before analysis, the dataset was checked for completeness and consistency across provinces and years. No missing-value imputation was performed, and the analysis proceeded using the available complete observations.

Table 1. Operational Variables and Definitions

Variabel	Operational Definition	Key Sub indicators	Scale
Civil Liberties (X ₁)	The level of protection for freedom of opinion, association, and expression.	Freedom of opinion, religion, association, freedom of the press.	0–100
Politics Rights (X ₂)	The degree of public participation in formal politics and public supervision.	Election participation, community representation, control over the government.	0–100
Institutions of Democracy (X ₃)	The effectiveness of local government institutions in upholding democratic principles.	The function of the DPRD, the effectiveness of local governments, the role of political parties.	0–100

Each province is represented by a three-dimensional feature vector, $X_i = [X_{i1}, X_{i2}, X_{i3}]$, comprising Civil Liberties, Political Rights, and Democratic Institutions. Because all three indicators use the same official 0–100 scale (table 1), the original scores were retained without additional normalization or standardization. Data preparation, descriptive examination, and preliminary statistical processing were conducted using JASP.

The RFC model is built on the concept of unsupervised random forest (Shi & Horvath, 2006), with the principle of forming random decision trees to produce a proximity matrix. The RFC procedure used 500 random trees to estimate similarities among provincial democracy profiles. The implementation followed an unsupervised random forest framework, in which forest-based proximity estimates were used to construct a dissimilarity matrix for subsequent clustering. JASP was used for data preparation and preliminary statistical analysis. The RFC and PAM computations were performed using [specify the actual software or computational implementation used], because the availability of these procedures depends on the software and its installed modules. The forest construction settings followed the actual model configuration used in the analysis.

The RFC procedure used 500 random trees to estimate similarities among provincial democracy profiles. The implementation followed an unsupervised random forest framework, in which forest-based proximity estimates were used to construct a dissimilarity matrix for subsequent clustering. JASP was used for data preparation and

preliminary statistical analysis. The RFC and PAM computations were performed using [specify the actual software or computational implementation used], because the availability of these procedures depends on the software and its installed modules. The forest construction settings followed the actual model configuration used in the analysis.

The interprovincial proximity matrix was calculated as follows:

$$P_{ij} = \frac{1}{K} \sum_{k=1}^K \mathbb{I}(T_k(X_i) = T_k(X_j)) \quad (1)$$

where P_{ij} denotes the proximity between provinces (i) and (j), (K) represents the number of trees, and $(T_k(X_i))$ represents the terminal node reached by province (i) in tree (k). The indicator function equals 1 when two provinces occupy the same terminal node and 0 otherwise. With 500 trees, the proximity value represents the proportion of trees in which two provincial observations share a terminal node.

The proximity matrix was converted into a dissimilarity matrix using the following expression:

$$D_{ij} = 1 - \frac{1}{K} \sum_{k=1}^K \mathbb{I}(T_k(i) = T_k(j)) \quad (2)$$

where D_{ij} denotes the dissimilarity between provinces (i) and (j). Values closer to zero indicate greater similarity, whereas higher values indicate greater dissimilarity between provincial democracy profiles.

This matrix (D) was then used as input for the clustering algorithm, Partitioning Around Medoids (PAM), to identify groups of provinces with similar democratic profiles.

PAM was applied to the RFC-derived dissimilarity matrix. Candidate cluster solutions were compared for different values of (k), and the final number of clusters was determined using the minimum BIC value, following the calculation procedure implemented in the analysis. The BIC calculation and the range of candidate cluster numbers should be reported according to the actual computational procedure used.

The analysis steps include pre-processing of data through normalization and standardization of IDI scores; The implementation of the RFC with the formation of 500 random trees to obtain the dissimilarity matrix between provinces; Determination of the optimal number of clusters based on the minimum BIC value; Evaluation of model performance using R^2 , Silhouette score, Calinski–Harabasz index, Dunn index, and Entropy; The visualization of the results uses the t-distributed Stochastic Neighbor Embedding (t-SNE) algorithm to map the relative positions

between provinces in two-dimensional space (Gosiewska & Biecek, 2020).

The evaluation metrics were used to assess complementary aspects of the resulting cluster structure. The Silhouette score evaluates within-cluster cohesion relative to separation from other clusters. The Calinski–Harabasz index assesses between-cluster dispersion relative to within-cluster dispersion, while the Dunn index evaluates cluster separation relative to cluster diameter. Entropy was used to assess the distribution of observations across clusters according to the definition adopted in the analysis. The (R^2) statistic was interpreted according to its specified calculation in the clustering context. Each metric was interpreted based on its definition and appropriate optimization direction.

The t-SNE algorithm was used exclusively to visualize the relative positions of provinces in two-dimensional space. Cluster membership was determined by the RFC-derived dissimilarity matrix and PAM rather than by the t-SNE coordinates. Data preparation and preliminary statistical examination were conducted using JASP, while the implementation of RFC, PAM, and the associated evaluation procedures followed the computational tools actually used to generate the results.

The justification for using RFCs is that this method does not assume the linearity of relationships between variables, it is more accurate in capturing complex interactions for governance and democracy analysis. To facilitate reproducibility, the study reports the analytical unit, the aggregation of annual IDI scores, the original indicator scales, the number of random trees, the proximity and dissimilarity calculations, the clustering procedure, and the evaluation criteria. The software and computational implementation used for RFC and PAM, together with the relevant parameter settings and random seed where available, should correspond to the procedures used to produce the reported results.

Results

This study analyzed IDI data from 34 provinces for the period 2021–2024, using three main dimensions: civil liberties, political rights, and democratic institutions. The RFC analysis was used to find patterns of interprovincial grouping empirically. The selection of the optimal number of clusters was determined through a comparison of BIC values and within-cluster sum of squares (WCSS) using the Elbow Method approach.

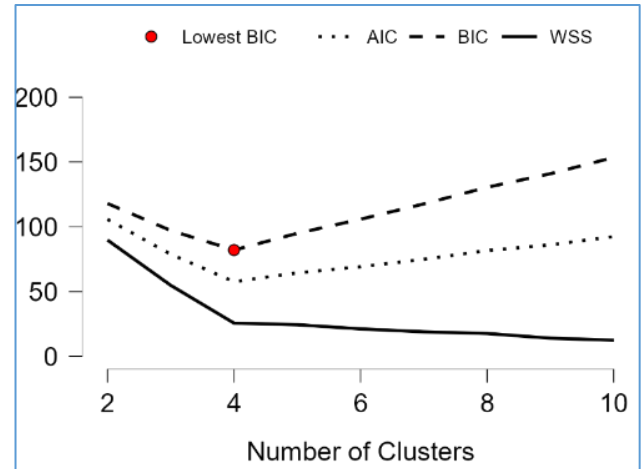


Figure 1. Elbow Method Plot for Determination of Optimal Cluster Number 2021–2024

Figure 1 shows the elbow point at $K = 4$, where the decline in the WCSS value begins to slope, indicating that the four-cluster model is the most efficient representation of the variation in democratic data between provinces. The determination of the optimal number of clusters by the elbow approach is widely used in unsupervised analysis because it shows a balance between model complexity and classification accuracy (Rousseeuw, 1987). In practice, the combination of the elbow method with evaluation metrics such as silhouette coefficient and BIC results in more empirically robust results (Antúnez et al., 2013). The model with four clusters was chosen because it provides the best balance between complexity and precision, as shown in Table 2.

Table 2. Summary of the RFC Model and Descriptive Statistics and Validity of IDI Random RFC in 2021–2024

Summary of the RFC Model of the IDI					
Number of Clusters	N	R^2	AIC	BIC	Silhouette
4	34	0.806	57.65	82.08	0.31
Descriptive Statistics and Validity of IDI RFC Clusters					
Cluster	Size (n)	Proportion of Heterogeneity in a Cluster	Within Sum of Squares	Silhouette	
1	14	0.321	8.242	0.295	
2	8	0.231	5.932	0.21	
3	8	0.224	5.737	0.37	
4	4	0.224	5.743	0.45	

Source: The results of the RFC analysis are based on BPS-IDI data (2021–2024).

Note: Between Sum of Squares = 106.35; Total Sum of Squares = 132.

The RFC model with $K = 4$ was selected as the optimal solution based on statistical performance: $R^2 = 0.806$ (explaining 80.6% of the data variation), lowest BIC (82.08), and silhouette score = 0.310, which indicates adequate internal cohesion and separation between clusters according to multivariate social analysis standards (Cole et al., 2021). RFC leverages the proximity matrix of Random Forest to identify latent structures in complex and non-linear social data—an advantage over classical regression (Waggoner, 2020). This unsupervised ensemble learning approach has proven to be stable in multidimensional data and has been used to analyze spatial inequalities in governance and democracy (Benati et al., 2025). Thus, the grouping into four provincial clusters—from weak to mature democracies—is a valid empirical representation of the complexity and spatial inequality of democracy in Indonesia, 2021–2024, in line with the trend of computational social research that combines machine learning and evidence-based spatial analysis (Little & Meng, 2024).

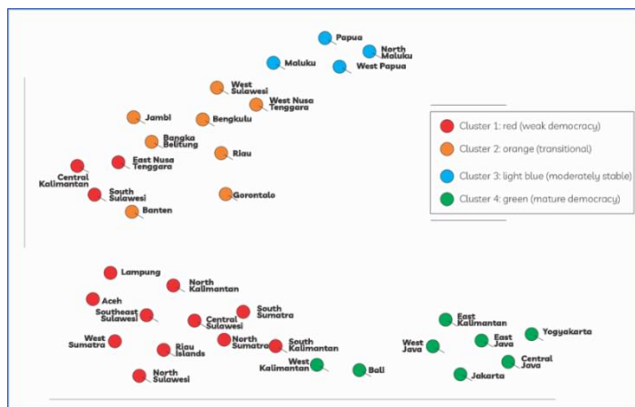


Figure 2. Visualization of Clustering with t-SNE Plot (IDI 2021–2024)

Table 2 shows the descriptive statistics and the validity of the resulting clusters. All four clusters showed size variation and internal heterogeneity, with the highest silhouette value in cluster 4 (0.45) indicating the strongest internal cohesion. The values of the Calinski–Harabasz index = 41.453 and the Dunn index = 0.197 indicate a stable and moderately separated cluster structure, while Entropy = 1.298 indicates a distribution between clusters that is still balanced. These results are consistent with the finding that the Calinski–Harabasz and Dunn indexes are able to detect relatively clear structures even in complex social data (Yang et al., 2021). RFC is an unsupervised ensemble learning approach that utilizes proximity matrices to group observations without labels, excels at capturing non-linear relationships between social and spatial variables and is more resistant to outliers than traditional methods such as k-means (Hastie et

al., 2009). The validity of the RFC model with four clusters was reinforced by a moderate positive correlation between metrics (Pearson's $\gamma = 0.494$), in accordance with a multicriteria validation approach in social cluster analysis (Renchunzi & Pratama, 2022) and ensemble-based validation methods to test resilience to random variation and spatial structure (Shi et al., 2023). By meeting internationally recognized multidimensional evaluation criteria (Ash et al., 2024), this model results in four typologies of regional democracy in Indonesia, from stable to low-participation democratic provinces, that form the basis for further analysis and evidence-based governance policy formulation.

The image shows the results of visualization of RFC for 34 provinces in Indonesia in the 2021–2024 period. Each color represents a group of provinces with similar characteristics of the Democracy Index. Cluster 1 ($n = 14$), had the highest heterogeneity (0.321), covering most provinces in Sumatra, Sulawesi, and central Kalimantan, which showed a relatively low level of democracy with challenges in terms of political participation and the effectiveness of local government institutions. Cluster 2 ($n = 8$) is characterized by the lowest silhouette (0.210), consisting of provinces such as Bengkulu, Jambi, and West Nusa Tenggara which are in the transition stage towards democratic consolidation, with achievements varying between indicators. Cluster 3 ($n = 8$), includes provinces on the islands of Java and Bali, such as Central Java, Yogyakarta, and DKI Jakarta, which shows democratic stability and better quality of public governance (silhouette 0.370). Meanwhile, Cluster 4 ($n = 4$), being the most stable group (silhouette 0.450), consists of provinces in the eastern region such as Papua, Maluku, and North Maluku, which shows relatively high democratic performance in the local context, despite complex socio-political dynamics. Overall, the pattern of cluster distribution shows the spatial inequality of democracy between regions of Indonesia, where the western and southern parts tend to be more institutionally established, while the central and eastern regions show greater variation in democratic consolidation.

Discussion

The results of the RFC on the IDI 2021–2024 reveal four provincial clusters that reflect the spatial inequalities of democracy, with Cluster 4 showing high democratic performance, characterized by effective institutions, civil liberties, and strong participation, according to the concept of democratic resilience (Levitsky & Way, 2020) and resilience to authoritarianism (Buben & Kouba, 2020), while Clusters 1 and 2 reflect weak democracy due to low

bureaucratic capacity, participation, and public trust, a common phenomenon in the context of spatial democratic inequality (Hömann & Nugent, 2022). This inequality is influenced by socio-economic factors, urbanization, and the quality of regional governance (Hesstvedt & Christensen, 2023), confirms the hypothesis that Indonesian democracy develops unevenly due to the territorial unevenness of democracy (Dosek, 2025). These findings confirm that the consolidation of national democracy depends on strengthening local democracy, and that this spatial inequality risks exacerbating polarization and erosion of public trust (Meza, 2015), thus requiring evidence-based policies to strengthen democratic resilience across regions.

Theoretically, this research expands the understanding of the concept of local democratic resilience, which is the ability of local democratic systems to maintain their stability in the midst of social, political, and institutional pressures. This concept is even more relevant in the contemporary era, when many democracies are under pressure due to governance inequality and weakening public participation. Comparative literature shows that the resilience of local democracy is highly dependent on institutional capacity, local government accountability, and the ability of civil society to resist the penetration of authoritarianism (Jacobsohn & Schor, 2018). In this context, resilience is not simply the absence of setbacks, but the ability to adapt to social and political shocks by strengthening participation and transparency at the local level (Whitehead, 2021). These findings are also in line with differential democracy theory, which emphasizes that variations in democratic performance are determined not only by national factors, but also by regional institutional capacity and civil society dynamics (Hidayat, 2018). Cross-border research shows that regions with adaptive institutions, independent media, and deliberative political cultures are better able to maintain the quality of democracy than regions with weak institutions (Levitsky & Way, 2020). Therefore, the spatial analysis of Indonesian democracy through the RFC approach makes an important empirical contribution to this theory, as it shows how interprovincial inequalities in political and social capacity affect the quality of local democracy. Furthermore, the results of this study also resonate with the global phenomenon of democratic backsliding, namely the decline in the quality of democracy in various developing countries due to increasing populism, erosion of public trust, and weak political institutions (Bhandari, 2023; Kelemen & McNamara, 2022; Lai et al., 2020). Empirically, backsliding often appears first at the local level, when regional elites use instruments of power to reduce the space for opposition and civil liberties (Boese et al., 2021). Thus,

spatial mapping of democracy through machine learning-based clustering can be an important tool for early detection of the dynamics of democratic decline. In terms of methodology, this study contributes to the machine learning-based literature of comparative democracy, because it is able to identify non-linear structures and latent patterns between provinces that are difficult to capture through conventional statistical analysis. Unsupervised learning approaches such as RFC allow for the exploration of complex democratic data without the assumption of linearity, as well as uncover latent relationships between public participation, institutional effectiveness, and civil liberties (Rifa'i & Abdullah, 2025). In the context of comparative politics, the application of machine learning has been shown to increase inferential validity and enable the discovery of new patterns that were previously hidden (Crawford & Paglen, 2021; Lee, 2025). In addition, the cutting-edge literature confirms the importance of integrating spatial analytics and ensemble learning to understand the variation of democracy across regions, as both are able to capture spatial effects and interactions between non-independent political units (Hadžić, 2022; Ikenberry & Gleditsch, 2003). Thus, the RFC approach not only makes an empirical contribution to the mapping of Indonesian democracy, but also expands the methodological horizon of the study of democracy governance and resilience at the subnational level. This study shows that strengthening the resilience of local democracy must start from increasing the capacity of local institutions, community political education, and the use of data-based analysis to detect spatial patterns of democratic inequality. Thus, these findings not only have theoretical implications for the study of differential and backsliding democracy, but also practical value in designing more adaptive and evidence-based decentralization policies in Indonesia.

Operationally, the results of this study provide an empirical basis for the formulation of evidence-based policies in strengthening regional democracy. This data-driven approach is important to ensure that government interventions are efficiently directed according to the needs of each regional democracy cluster. Provinces in Clusters 1 and 2, which show weaknesses in the capacity of political institutions and public participation, need to be a priority for intervention through improving the institutional quality of the District House of Representatives (DPRD), citizen political literacy, and strengthening public supervisory institutions. This approach is in line with the findings of (Zhang et al., 2022), which affirm the importance of improving political literacy and legislative accountability to improve the quality of subnational democracy. In addition,

a study by (Hjermitslev, 2023) shows that the capacity of local legislative institutions has a significant effect on governance stability and adaptability to political pressures. Therefore, increasing the capacity of the DPRD in weak provinces is a strategic step to strengthen the resilience of local democracy. On the other hand, community political literacy programs have also been shown to increase civic engagement and reduce participatory gaps between regions (Balintongog & Narido, 2025; Pujiati & Badariah, 2025). Provinces in Cluster 3 can be positioned as transition zones that require participatory policy support to stabilize the dynamics of local democracy. A collaborative approach between local governments, civil society organizations, and educational institutions can strengthen social capital and expand public participation mechanisms (Polizzi, 2025). In this context, inclusive participatory policies have been shown to increase the legitimacy and effectiveness of local government, especially in areas with high social diversity (Haesevoets et al., 2024; Simpson et al., 2025). Meanwhile, Cluster 4, which reflects high democratic performance, can be used as a model of best practices to be replicated by other provinces. A study (Flinders, 2022) confirms that transparent and collaborative governance practices in the regions contribute to increased public trust and policy effectiveness. Thus, the central government can develop a knowledge-sharing framework to disseminate successful democratic governance practices in these regions. Furthermore, the application of adaptive governance principles is also needed to adapt policy strategies to changing local political dynamics. According to (Datta & Chaffin, 2022), adaptive policy design allows for the strengthening of local institutions through feedback mechanisms and continuous policy learning. This approach is in line with the idea of place-sensitive governance, where strategies for strengthening democracy must take into account the unique local social, economic, and cultural context (Panday, 2025; Yıldız et al., 2025). The results of this study confirm that effective policies to strengthen regional democracy must be differential, adaptive, and evidence-based. The government can use the results of this clustering to design a roadmap for strengthening provincial democracy that is directed with data-based indicators and periodic evaluation using machine learning methods to monitor the effectiveness of policies over time. This kind of approach has been shown to improve governance accountability and policy effectiveness in various countries (Papadopoulos, 2023; Porumbescu et al., 2025). Thus, the results of this study not only provide an academic contribution to the literature on local democratic governance, but also practical recommendations for policy makers in building a more inclusive, resilient, and data-based democracy.

The application of the RFC in democracy analysis has global strategic value as an empirical classification tool as well as an early-warning system against democratic erosion and institutional decay, thanks to its ability to uncover non-linear patterns in multidimensional socio-political data (Klüser & Breunig, 2023). This machine learning-based approach is aligned with SDG-16 on peace, justice, and strong institutions, as it allows for the identification of vulnerable areas before a crisis of legitimacy occurs (Glass & Newig, 2019), while supporting evidence-based policies in public governance (Pedersen & Favero, 2020). RFCs excel in spatial mapping of governance quality because they are able to capture latent relationships between regional inequalities and democratic resilience that are not seen in conventional regression (Sadeghi et al., 2023), strengthening predictive governance models based on spatial simulations and historical trends (Wu et al., 2025). Thus, the application of the RFC to the IDI is not only a methodological contribution to the study of comparative democracy, but also an operational tool to strengthen institutional resilience and support the global agenda of SDG-16 (UNDP, 2024).

Conclusion

This research shows that democracy in Indonesia still faces spatial inequality between provinces during the 2021–2024 period. Through the application of Random Forest Clustering, four main clusters were found that illustrate variations in the level of regional democracy. The first and second clusters consist of provinces with low political participation and limited institutional capacity, while the third cluster shows fluctuating democratic performance that still needs stabilization. In contrast, the fourth cluster reflects a region with mature democracy, characterized by active citizen participation, strong civil liberties, and responsive governance. The R^2 value of 0.806, the Silhouette score of 0.310, and the Calinski–Harabasz index of 41.453 indicate the robustness of the cluster structure and the validity of the model. This result confirms that regional democratic inequality is a national strategic issue, because it has a direct impact on the effectiveness of decentralization and the balance of regional-based governance. In a global context, these findings also reflect the challenges to democratic resilience faced by many developing countries amid increasing populism pressures and weak political institutions.

In the future, this research can be further developed by expanding the analysis variables that include social, economic, and digitalization aspects of government to understand the structural determinants of regional

democracy more comprehensively. In addition, the use of longitudinal data until post-2024 will allow an evaluation of the temporal dynamics of democracy between provinces. A comparative approach between countries in the Southeast Asian region is also recommended to test the suitability of clustering models in regional contexts. The integration of GIS-based spatial analysis can strengthen understanding of the geographical relationship between the quality of democracy and regional development. Thus, the application of machine learning such as RFC is not only an empirical analysis tool, but also a strategic instrument to support evidence-based policy formulation in strengthening local democracy and national institutional resilience.

References

- Adam, C., Hurka, S., Knill, C., & Steinebach, Y. (2022). On democratic intelligence and failure: The vice and virtue of incrementalism under political fragmentation and policy accumulation. *Governance*, 35(2). <https://doi.org/10.1111/gove.12595>
- Antúnez, E., Marfil, R., Bandera, J. P., & Bandera, A. (2013). Part-based object detection into a hierarchy of image segmentations combining color and topology. *Pattern Recognition Letters*, 34(7). <https://doi.org/10.1016/j.patrec.2013.01.012>
- Ash, E., Gauthier, G., & Widmer, P. (2024). Semantics Capture Political and Economic Narratives. *Political Analysis*, 32(1), 115–132. <https://doi.org/10.1017/pan.2023.8>
- Balintongog, A., & Narido, N. (2025). Political Awareness Among Residents of Select Barangays in Claveria, Misamis Oriental. *International Journal on Science and Technology*, 16(3). <https://doi.org/10.71097/ijst.v16.i3.7815>
- Bartels, L. M., Daxecker, U. E., Hyde, S. D., Lindberg, S. I., & Nooruddin, I. (2023). The Forum: Global Challenges to Democracy? Perspectives on Democratic Backsliding. *International Studies Review*, 25(2). <https://doi.org/10.1093/isr/viad019>
- Benati, S., Bon, M., & Nardi, F. (2025). Exploring the predictors of the populist vote using random forests. *Quality and Quantity*, 59(Suppl 2). <https://doi.org/10.1007/s11135-025-02064-3>
- Bhandari, A. (2023). Social, Formal, and Political Determinants of Trade Under Weak Rule of Law: Experimental Evidence from Senegalese Firms. *Comparative Political Studies*, 56(2). <https://doi.org/10.1177/00104140221089648>
- Boese, V. A., Edgell, A. B., Hellmeier, S., Maerz, S. F., & Lindberg, S. I. (2021). How democracies prevail: democratic resilience as a two-stage process. *Democratization*, 28(5). <https://doi.org/10.1080/13510347.2021.1891413>
- Boese, V. A., Gates, S., Knutsen, C. H., Nygård, H. M., & Strand, H. (2022). Patterns of Democracy over Space and Time. *International Studies Quarterly*, 66(3). <https://doi.org/10.1093/isq/sqac041>
- BPS. (2025). Indeks Demokrasi Indonesia menurut Provinsi, 2024. Badan Pusat Statistik. <https://www.bps.go.id/id/statistics-table/2/MjE1OSMy/-metode-baru-indeks-demokrasi-indonesia-menurut-provinsi.html>
- Buben, R., & Kouba, K. (2020). Nicaragua in 2019: The surprising resilience of authoritarianism in the aftermath of regime crisis*1. *Revista de Ciencia Política*, 40(2). <https://doi.org/10.4067/s0718-090x2020005000114>
- Cole, C. L., Sengupta, S., Rossetti (née Collins), S., Vawdrey, D. K., Halaas, M., Maddox, T. M., Gordon, G., Dave, T., Payne, P. R. O., Williams, A. E., & Estrin, D. (2021). Ten principles for data sharing and commercialization. *Journal of the American Medical Informatics Association*, 28(3), 646–649. <https://doi.org/10.1093/jamia/ocaa260>
- Crawford, K., & Paglen, T. (2021). Excavating AI: the politics of images in machine learning training sets. *AI and Society*, 36(4). <https://doi.org/10.1007/s00146-021-01162-8>
- Datta, A. W., & Chaffin, B. C. (2022). Evolving adaptive governance: challenging assumptions through an examination of fisheries law in Solomon Islands. *Ecology and Society*, 27(2). <https://doi.org/10.5751/ES-13251-270230>
- Dosek, T. (2025). The Persistence of Local Caudillos in Latin America: The Case of Vicky Barahona in Santiago de Chile (2000–2016). *Studies in Comparative International Development*. <https://doi.org/10.1007/s12116-025-09467-1>
- Flinders, M. (2022). Measuring accountability in public governance regimes. *EllenRockCambridge University Press*, New York, NY, 2020. 256 pp. \$110 (cloth). *Governance*, 35(1). <https://doi.org/10.1111/gove.12654>
- Fukuyama, F., Dann, C., & Magaloni, B. (2025). DELIVERING FOR DEMOCRACY: WHY RESULTS MATTER. *Journal of Democracy*, 36(2). <https://doi.org/10.1353/jod.2025.a954557>
- Gagnon, J. P., & Fleuss, D. (2020). The case for extending measures of democracy in the world “Beneath”, “Above”, and “Outside” the national level. *Political Geography*, 83. <https://doi.org/10.1016/j.polgeo.2020.102276>
- Glass, L.-M., & Newig, J. (2019). Governance for achieving the Sustainable Development Goals: How important are participation, policy coherence, reflexivity, adaptation and democratic institutions? *Earth System Governance*, 2. <https://doi.org/10.1016/j.esg.2019.100031>
- Gosiewska, A., & Biecek, P. (2020). Do Not Trust Additive Explanations.
- Grubman, N. (2022). Transition Arrested. *Journal of Democracy*, 33(1). <https://doi.org/10.1353/jod.2022.0001>
- Hadžić, F. (2022). Social Scientific Centered Issues Status: In Production Islam and Europeanism; cultural, religious, and security critique of the contemporary neo-nationalism and neo-radicalism. *Social Scientific Centered Issues Journal*, 4(1).
- Haesevoets, T., Roets, A., Steyvers, K., Verschuere, B., & Wauters, B. (2024). Towards a multifaceted measure of perceived legitimacy of participatory governance. *Governance*, 37(3). <https://doi.org/10.1111/gove.12800>
- Harbers, I., Bartman, J., & van Wingerden, E. (2019). Conceptualizing and measuring subnational democracy across Indian states. *Democratization*, 26(7). <https://doi.org/10.1080/13510347.2019.1606797>

- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer New York. <https://doi.org/10.1007/978-0-387-84858-7>
- Hellmeier, S., Cole, R., Grah, S., Kolvani, P., Lachapelle, J., Lührmann, A., Maerz, S. F., Pillai, S., & Lindberg, S. I. (2021). State of the world 2020: autocratization turns viral. *Democratization*, 28(6). <https://doi.org/10.1080/13510347.2021.1922390>
- Hesstvedt, S., & Christensen, J. (2023). Political and administrative control of expert groups—A mixed-methods study. *Governance*, 36(2), 337–357. <https://doi.org/10.1111/gove.12599>
- Hidayat, R. (2018). Democratic Citizenship of Teacher Movement in Indonesia Post-Soeharto: Between Democratic Citizenship and Civic Engagement. *Jurnal Ilmiah Peuradeun*, 6(3). <https://doi.org/10.26811/peuradeun.v6i3.236>
- Hjermitslev, I. B. (2023). Collaboration or competition? Experimental evidence for coalition heuristics. *European Journal of Political Research*, 62(1). <https://doi.org/10.1111/1475-6765.12537>
- Höhm, D., & Nugent, M. (2022). Male MPs, electoral vulnerability and the substantive representation of women's interests. *European Journal of Political Research*, 61(3), 762–782. <https://doi.org/10.1111/1475-6765.12472>
- Ikenberry, G. J., & Gleditsch, K. S. (2003). All International Politics Is Local: The Diffusion of Conflict, Integration, and Democratization. *Foreign Affairs*, 82(1). <https://doi.org/10.2307/20033441>
- Jacobsohn, G., & Schor, M. (2018). Comparative constitutional theory. In *Comparative Constitutional Theory*. <https://doi.org/10.17159/1727-3781/2019/v22i0a6173>
- Kelemen, R. D., & McNamara, K. R. (2022). State-building and the European Union: Markets, War, and Europe's Uneven Political Development. *Comparative Political Studies*, 55(6). <https://doi.org/10.1177/00104140211047393>
- Klüser, K. J., & Breunig, C. (2023). Ministerial policy dominance in parliamentary democracies. *European Journal of Political Research*, 62(2). <https://doi.org/10.1111/1475-6765.12545>
- Lai, K. M. Y., Saffar, W., Zhu, X. (Kevin), & Liu, Y. (2020). Political institutions, stock market liquidity and firm dividend policy: Some international evidence. *Journal of Contemporary Accounting and Economics*, 16(1). <https://doi.org/10.1016/j.jcae.2019.100180>
- Lee, F. (2025). The practices and politics of machine learning: a field guide for analyzing artificial intelligence. *AI and Society*, 40(8). <https://doi.org/10.1007/s00146-025-02430-7>
- Levitsky, S., & Way, L. (2020). The new competitive authoritarianism. *Journal of Democracy*, 31(1). <https://doi.org/10.1353/jod.2020.0004>
- Little, A. T., & Meng, A. (2024). Measuring Democratic Backsliding. *PS - Political Science and Politics*, 57(2). <https://doi.org/10.1017/S104909652300063X>
- Liu, W., Dridi, M., Fei, H., & El Hassani, A. H. (2021). Hybrid metaheuristics for solving a home health care routing and scheduling problem with time windows, synchronized visits and lunch breaks. *Expert Systems with Applications*, 183. <https://doi.org/10.1016/j.eswa.2021.115307>
- Meza, O. D. (2015). Local Governments, Democracy, and Inequality. *State and Local Government Review*, 47(4). <https://doi.org/10.1177/0160323x15627852>
- OECD. (2022). *Regional Governance in OECD Countries*. OECD Publishing. <https://doi.org/10.1787/4d7c6483-en>
- Panday, P. K. (2025). Deficits in Democratic Governance in Developing Countries: The Bangladesh Scenario. *Journal of Asian and African Studies*, 60(4). <https://doi.org/10.1177/00219096231215694>
- Papadopoulos, Y. (2023). Understanding Accountability in Democratic Governance. In *Understanding Accountability in Democratic Governance*. <https://doi.org/10.1017/9781108973823>
- Pavel, M. (2025). *Quality of Democracy at Regional Level in the V4 Countries*. Routledge. <https://doi.org/10.4324/9781003591894>
- Pedersen, M. J., & Favero, N. (2020). Social Distancing during the COVID-19 Pandemic: Who Are the Present and Future Noncompliers? In *Public Administration Review* (Vol. 80, Issue 5). <https://doi.org/10.1111/puar.13240>
- Polizzi, G. (2025). Digital literacy and strategic (dis)engagement: examining how functional and critical digital literacy shapes participation. *Information Communication and Society*, 28(14). <https://doi.org/10.1080/1369118X.2025.2452282>
- Porumbescu, G., Young, M., & Yang, K. (2025). Introduction to Special Issue on Generative AI and Public Administration: Performance, Accountability, and Democratic Governance. In *Public Performance and Management Review* (Vol. 48, Issue 6). <https://doi.org/10.1080/15309576.2025.2570315>
- Pujiati, P., & Badariah, N. (2025). Criticism of Digital Native Students in Responding to Local Digital Democracy: Kritisalitas Mahasiswa Digital Native dalam Menyikapi Demokrasi Digital Lokal. In *Indonesian Journal of Innovation Studies* (Vol. 26, Issue 3).
- Renchunzi, X., & Pratama, M. (2022). Automatic online multi-source domain adaptation. *Information Sciences*, 582, 480–494. <https://doi.org/10.1016/j.ins.2021.09.031>
- Rifa'i, I. J., & Abdullah, A. (2025). The Role of Law in Strengthening Democratic Governance. *Journal of Law and Social Politics*, 3(2). <https://doi.org/10.46799/jlsp.v3i1.62>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20(C). [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Sadeghi, M., Karimi, M., Rabiei-Dastjerdi, H., & Sarkar, D. (2023). Adaptive weighted least squares (AWLS): A new vector-based model to improve urban population estimation at small-area scale using morphology and attractiveness criteria. *Applied Geography*, 158. <https://doi.org/10.1016/j.apgeog.2023.103050>
- Schneider, C. J. (2018). The Domestic Politics of International Cooperation. In *Oxford Research Encyclopedia of Politics*. <https://doi.org/10.1093/acrefore/9780190228637.013.615>
- Shi, Y., Wang, B., Yu, Y., Tang, X., Huang, C., & Dong, J. (2023). Robust anomaly detection for multivariate time series through temporal GCNs and attention-based VAE. *Knowledge-Based Systems*, 275, 110725. <https://doi.org/10.1016/j.knosys.2023.110725>

- Simpson, C., Beck, A., & Campbell, L. (2025). The review and development of professional standards through the lens of Democratic Anchorage points. *British Educational Research Journal*, 51(2). <https://doi.org/10.1002/berj.4090>
- UNDP. (2024). Reimagining cooperation in a polarized world.
- Waggoner, P. D. (2020). *Unsupervised Machine Learning for Clustering in Political and Social Research*. Cambridge University Press. <https://doi.org/10.1017/9781108883955>
- Whitehead, L. (2021). Latin America Erupts: The Danger of Democratic Delinquency. *Journal of Democracy*, 32(3), 78–93. <https://doi.org/10.1353/jod.2021.0035>
- Wu, B., Wang, S., Zou, Y., Chen, Y., Luo, Y., Chen, Y., & Li, Y. (2025). Multi-scenario simulation of land use spatial conflicts: A spatially explicit model integrating conflict risk and ecological impact. *Environment, Development and Sustainability*. <https://doi.org/10.1007/s10668-025-06329-7>
- Yang, D. H., Zhou, X., Wang, X. Y., & Huang, J. P. (2021). Micro-earthquake source depth detection using machine learning techniques. *Information Sciences*, 544. <https://doi.org/10.1016/j.ins.2020.07.045>
- Yıldız, B. F., Alola, A. A., & Wong, W. K. (2025). Socioeconomic Development Aspects of Democratic Governance across Selected Countries. *Democracy and Security*, 21(4). <https://doi.org/10.1080/17419166.2023.2178422>
- Zhang, D., Pee, L. G., Pan, S. L., & Cui, L. (2022). Big data analytics, resource orchestration, and digital sustainability: A case study of smart city development. *Government Information Quarterly*, 39(1). <https://doi.org/10.1016/j.giq.2021.101626>