FWU Journal of Social Sciences, Summer 2026, Vol.20, No.2, 164-185
DOI: http://doi.org/10.51709/19951272/Summer2026/12

Analyzing Multidimensional Poverty in Roma Settlements: A WEFE Nexus and Machine Learning Approach

Tabish Nawab
Ibn Haldun University, Istanbul, Türkiye

Snežana Radukić
University of Niš, Serbia

This study aims to analyze multidimensional poverty determinants within Roma settlements in Serbia, North Macedonia, and Montenegro using logistic regression and machine learning to identify the socioeconomic and resource-based components influencing poverty within the WEFE (Water Energy Food Ecosystems) Nexus. Reliable lighting, sanitation maintenance, safe water access, consistent water supply, and energy for cooking all play a critical role in poverty alleviation. Our key findings align with SDG 1 (No Poverty), SDG 6 (Clean Water and Sanitation), and SDG 7 (Affordable and Clean Energy), underscoring the global significance and relevance of our research. Random Forest and Extra Trees perform very well when compared to logistic regression by capturing highly variable interactions that may be missed by logistic regression. Results with country-specific emphases are presented, such as digital access in Montenegro and household size in North Macedonia, to illustrate the adaptability of the WEFE framework to different regional contexts. The results advocate for resource-driven integrated policies to improve people's access to important utilities, financial inclusion, and digital connectivity to build social sustainability and resilience. The study supports NexusNet's plan to lead SDG-aligned poverty reduction in all sectors in the Western Balkans by focusing on WEFE resources and socio-economic supports .

***Keywords**:* WEFE nexus, multidimensional poverty, machine learning, logistic regression, SDGs, roma settlements, Western Balkan, social sustainability

Multidimensional poverty is a deeply embedded problem in the Western Balkans where many vulnerable groups, in particular the Roma communities[1] in the region, suffer from very widespread deprivations in several dimensions (e.g., lack of access to water, energy, and food) (Morar et al., 2004). Such intersecting deprivations are often left out of traditional poverty measures, which are based on income. For this, a multidimensional approach taking into account the social, economic, and environmental facets is needed. More recent advancements in the Water, Energy, Food, and Ecosystems (WEFE) Nexus paradigm and machine learning provide more promising possibilities for enhanced poverty analysis and intervention strategy. This study, which aims to develop sustainable solutions for poverty alleviation around Roma settlements located in Serbia, North Macedonia and Montenegro, is firmly aligned with the United Nations Sustainable Development Goals (SDGs). This alignment, particularly with SDG 1 (No Poverty), SDG 6 (Clean, Water and Sanitation), SDG 7 (Affordable and Clean Energy), SDG 10 (Reduced Inequalities), and SDG 11 (Sustainable Cities and Communities), underscores the relevance and importance of the study.

Correspondence concerning this article should be addressed to Tabish Nawab, Department of Economics, Ibn Haldun University, Istanbul, Türkiye. Email: tabish.nawab@stu.ihu.edu.tr

[1] The Roma community, also known as Gypsies, are a traditionally nomadic ethnic group originating from India who migrated to Europe between the 5th and 10th centuries A.D. Distinct cultural practices and languages characterize them and have faced historical marginalization.

The WEFE Nexus framework, a comprehensive approach, serves as a starting point to understand how interdependencies between water, energy, food, and natural ecosystems create poverty and constrain resource access. It integrates resource management to promote efficiency and sustainability, thereby directly supporting SDG 6 on access to essential resources, SDG 7 and 15. The structured nature of the WEFE Nexus will be of benefit to marginalized Roma communities who traditionally face resource scarcity and inequities, which can be addressed within the framework of the WEFE Nexus (Ivanov et al., 2015; El-Meligi., et al., 2024). Focusing on the equitable distribution of resources, this framework is a key player in reducing vulnerability to health, education, and livelihood, thereby ensuring the social sustainability of Roma settlements (Cheng, 2019; UNECE, 2022).

The Multidimensional Poverty Index (MPI) enhances the WEFE Nexus by introducing a complementary measure of poverty, focusing on deprivations in health, education, and living standards (Alkire & Foster, 2011). This is in line with SDG 1, which views poverty as a multidimensional issue necessitating comprehensive intervention strategies. The MPI's ability to reveal overlapping deprivations is particularly beneficial for the Roma settlements, which face a range of issues, including inadequate sanitation, substandard housing, and low educational levels, all of which contribute to the persistence of poverty (Ivanov et al., 2015; Ahmadi et al., 2024). However, it's important to note that the MPI, while powerful, is not fully equipped to account for the intricate nature of multi-layered poverty. This is where machine learning steps in, offering a potential solution to this complex issue.

Machine learning models can analyze complex, multidimensional datasets to reveal hidden patterns and determinants of poverty. This makes machine learning models highly attractive for development policy (Salas-Rojo & Rodríguez, 2022). These models capture the non-linearity and the interaction effects, improve poverty prediction, and uncover important poverty determinants (Lekobane & Samboma, 2024). The value of this analytical depth is especially relevant for WEFE Nexus-type interventions because machine learning models can prioritize poverty drivers within Roma communities, which can help inform, in turn, the more precise and effective use of resources. Robust predictive performance of ensemble models such as the Random Forest, Extreme Gradient Boosting, XG Boost) has been found to yield insights that harmonize with SDG 10 on resource access inequalities (Cheng, 2019).

This research takes the WEFE Nexus, machine learning, and MPI together and applies them to the analysis of poverty determinants in Roma settlements situated in Serbia, North Macedonia, and Montenegro. This research examines resource disparities to identify how poverty results from these disparities and how resource-based interventions aimed at poverty reduction can make a difference in Roma communities. Within the WEFE Nexus framework, we implement machine learning approaches that allow for data-driven SDG 1 strategies and emphasize the significance of equitable access to water, energy, and sanitation as in line with SDG 6 and SDG 7. In the end, this serves to support the implementation of sustainable, inclusive development pathways in addressing the socio-economic problems of marginalized populations in Western Balkans in line with SDG 11 and SDG 15.

Overall, this study presents a way to tackle multidimensional poverty using a new approach: the combination of the WEFE Nexus framework, MPI, and machine learning. Comprising relationships between resource dependencies and socio-economic factors, it offers a one-stop view of poverty. Finally, the study focuses on resource access inequalities and its deployment of advanced predictive analytics to provide actionable insights for policymakers and practitioners interested in reducing poverty and achieving sustainable development in Roma settlements.

Literature Review

Roma settlements across Serbia, North Macedonia, and Montenegro suffer from multidimensional poverty, which is alleviated through the use of integrated approaches. By focusing on communities facing resource inequalities employing the WEFE Nexus framework, MPI, and machine learning, this study lays out a solution. Together, these tools provide a structured approach to social sustainability and poverty reduction aligned with related SDGs

WEFE Nexus framework indicates the interconnectivity of water, energy, food, and ecosystems for better and more efficient usage on an integrated basis (Al-Saidi & Elagib, 2017; El-Meligi., et al., 2024). This provides a comprehensive approach to addressing multidimensional poverty for Roma communities, most of which are subject to considerable resource constraints. The Nexus framework closely supports SDG 6, 7, and 15 by supporting efficient resource use (Bazilian et al., 2011). Research has shown that the inefficiencies caused by the separate management of resources are overcome by WEFE-based frameworks (Ringler et al., 2013; Rasul, 2016), which lead to equitable distribution of resources that help mitigate poverty. The Nexus can contribute to reducing poverty in underserved areas through improved provision of key resources that are important to meet SDG 10 (Cheng, 2019; UNECE, 2022;). Through the Nexus framework, such socio-economic inequities that perpetuate the deprivation in Roma settlements are supported in resource allocation and decision-making as it addresses the community's inherent need for socio-economic sustainability (Rochovská, & Rusnáková, 2018; Orton et al., 2019; Ahmadi et al., 2024).

The Multidimensional Poverty Index (MPI) combines poverty with measures of deprivations in notions of health, education, and living standards (Alkire & Foster, 2011; Alkire & Santos, 2013). WEFE Nexus approach with poverty complexity and interconnectivity is completed by this approach (Mikula et al., 2018). MPI assessments in the Western Balkans document significant disparities, especially in Roma settlements, where communities suffer severe deprivation in sanitation, housing, and education (Matlovičová et al., 2022; Ahmadi et al., 2024; Rashid, & Kulub, 2026).

Through the integration of MPI and WEFE resources, these deprivations can be alleviated in a targeted manner, which contributes to UNECE's key environment infrastructure policies and SDGs 6 and 7 (UNECE, 2022). This framework aligned with SDG 1 because it uses resource-based, socially sustainable interventions to tackle poverty issues (El-Meligi., et al., 2024).

By processing complex, multidimensional data, machine learning not only identifies key poverty determinants but also extends the notion of symptom detection towards poverty. This enables the practical application of machine-learned poverty analysis (Jean et al., 2016; Lekobane et al., 2024). Ensemble models such as Random Forest, Extreme Gradient Boosting, and XG Boost, particularly useful for understanding Roma settlement poverty dynamics, are effective in uncovering non-linear relationships and interactions between poverty factors (Cheng, 2019; Hassan, 2024).

Using machine learning models, the WEFE Nexus framework sorts feature importance, which allows policymakers to use data-driven insights on what to prioritize with resources. Machine learning's predictive capabilities align with and support SDGs 6 and 7, where machine learning's predictive ability is directed to urgent needs such as the demonstration of 'access' to water, energy, and sanitation (Rasul, 2016; Salas-Rojo and Rodríguez, 2022; El-Meligi et al., 2024). SDG 10 is supported, as is resource access inequality among marginalized communities (Orton et al., 2019; Matlovičová et al., 2022; Lekobane et al., 2024).

Water, sanitation, energy, and food are essential resources for reducing poverty and social sustainability (Bazilian et al., 2011; Ringler et al., 2013; Trbojević et al., 2024). Better health and education and better economic outcomes, all of which are fundamental to breakdown cycles of deprivation in Roma settlements, are associated with better resource access (Orton et al., 2019; UNECE, 2022). By supporting SDG 1, enhanced resource access can increase productivity and well-being, which can be the part of poverty alleviation (Cheng, 2019).

Drawing from resource-based strategies, the WEFE Nexus framework advocates for policies and strategies that ensure resilience and social equity and promote equal opportunity (Al-Saidi & Elagib, 2017; El-Meligi et al., 2024). This study examines resource access disparities in Serbia, North Macedonia, and Montenegro and targets interventions that can directly alleviate local socio-economic circumstances to a degree while making progress towards aligned SDGs (Trbojević et al., 2024).

This study presents a holistic framing of poverty's socioeconomic dynamics using an integrated WEFE Nexus framework and machine learning. The Nexus builds upon integrated resource management to achieve SDGs and machine learning facilitates data-driven insights on poverty determinants (UNECE, 2022; El-Meligi et al., 2024). Machine learning helps identify resource gaps in Roma settlements, so targeted interventions can guide policymakers toward effective poverty alleviation strategies (Jean et al., 2016; Lekobane et al., 2024).

However, a conspicuous gap in the research exists on how the integrated WEFE Nexus, MPI, and machine learning approaches can achieve robust long-term poverty reduction as they relate to the particular socio-cultural needs of Roma communities. This gap, which urgently needs to be addressed, is hindering the creation of sustainable, context-based strategies that will effectively reduce poverty while also elevating Roma settlements to resilient social pathways toward social equity and inclusion.

Method

As a primary analytical tool, we used the Multidimensional Poverty Index (MPI) to measure multidimensional poverty across Roma settlements in Serbia, North Macedonia, and Montenegro in order to make accurate measurements. Using Alkire and Foster's (2011)[2] Framework, we are able to assess simultaneously poverty along multiple dimensions of education, health, and living standards using the MPI framework. To conduct this household-level analysis, we used a multiple indicator cluster survey (MICS)[3], which is a reliable cross-country dataset specifically designed to assess a wide range of health, education, and socioeconomic indicators. MICS data provide a consistent framework for the analysis of household-level deprivations among these Roma populations at a time when there are losses in the comparability of measurement across these populations in other datasets.

Data

The Multiple Indicator Cluster Survey (MICS) is the primary dataset used for this study, a globally recognized demo Figureic and health survey developed by UNICEF and the United Nations Development Program (UNDP). MICS gathers information on a range of indicators, including mortality, nutrition, maternal and child health, water and sanitation access, reproductive health, child development, and education, as well as on socioeconomic conditions and healthy well-being. MICS was initially conceived as a method to monitor the Millennium Development Goals (MDGs) but is now a key tracking instrument of progress towards the Sustainable Development Goals (SDGs). Its role in tracking progress towards the SDGs underscores its importance and makes it an ideal source from which to construct the MPI in Roma settlements, where multidimensional poverty is primarily driven by a lack of access to clean water, energy, sanitation, and educational opportunities.

Three MICS surveys, each with a substantial sample size, were conducted in Roma settlements to gain insights from marginalized communities. These surveys provide cross-nationally comparable measures of deprivation at the household level, including the extent of deprivation in Roma communities. In Serbia's

---

[2] Alkire, S., & Foster, J. (2011). *Counting and multidimensional poverty measurement*. Journal of Public Economics, 95(7-8), 476-487. doi: 10.1016/j.jpubeco.2010.11.006

[3] https://mics.unicef.org/

MICS 2019[4], a total of 1,934 households (8,329 people) participated, with an impressive response rate of 96.8 percent. The MICS 2018-19[5] in North Macedonia sampled 6,249 individuals in 1,584 households, with a response rate of 94.2 percent, while the MICS 2018[6] in Montenegro included 4,732 individuals in 1,165 households, with a response rate of 80.9 percent.

Although the MICS surveys provide an important and comparatively consistent basis for household-level analysis in Roma settlements, the sample composition still requires cautious interpretation. In particular, differences in the urban-rural distribution of surveyed households and variation in country-specific sample sizes may influence the precision and comparability of the estimates. For this reason, the findings should be interpreted primarily as applying to the surveyed Roma settlements included in the MICS samples rather than as automatically representative of all Roma communities in the Western Balkans. This consideration is especially important where settlement conditions, infrastructure access, and deprivation profiles differ between urban and rural locations.

Area of Study

The study focuses on Roma settlements within three Western Balkan countries: North Macedonia and Montenegro. Roma and a number of these nations share a history of similar socio-economic development, which is subject to study in order to arrive at the determinants of poverty in Roma communities. Systemic inequalities in resource access and social inclusion dominate Roma settlements in these countries, and as such, they are essential for understanding poverty and developing targeted interventions (Bieber, 2005; Rodríguez-Pose & Stermšek, 2015; World Bank, 2019).

A Machine Learning Model Selection and Validation

In addition to logistic regression, several supervised machine-learning models were evaluated to predict multidimensional poverty across Roma settlements in Serbia, North Macedonia, and Montenegro. The comparative modelling framework included Decision Tree, Support Vector Machine, Logistic Regression, Gradient Boosting, AdaBoost, XGBoost, LightGBM, Random Forest, and Extra Trees. Among these, Random Forest and Extra Trees were selected as the principal models for interpretation because they are particularly appropriate for datasets characterized by heterogeneous socio-economic and WEFE-related predictors, potential non-linear relationships, and interaction effects among variables. These ensemble tree-based approaches are also comparatively robust in handling complex classification tasks and provide a useful basis for identifying the relative importance of predictors in multidimensional poverty analysis.

Model performance was assessed using a cross-validation-based evaluation procedure in order to reduce the risk of overfitting and improve the reliability of model comparison. Hyperparameter tuning was undertaken to optimize model specifications and to identify the best-performing configurations for each algorithm. The final emphasis on Random Forest and Extra Trees was based not only on their strong predictive performance across the three country settings, but also on their relative stability and interpretive usefulness in the context of policy-oriented poverty analysis.

Operationalization of the WEFE Nexus in the Empirical Framework

The WEFE Nexus framework was operationalized in this study as the conceptual basis for identifying and interpreting key household resource conditions linked to multidimensional poverty. In practice, this meant that variables related to water, energy, and sanitation-related service conditions were selected as core

---

[4] Statistical Office of the Republic of Serbia and UNICEF. 2020. Serbia Multiple Indicator Cluster Survey and Serbia Roma Settlements Multiple Indicator Cluster Survey, 2019, Survey Findings Report. Belgrade, Serbia: Statistical Office of the Republic of Serbia and UNICEF.

[5] State Statistical Office and UNICEF. 2020. *2018-2019 North Macedonia Multiple Indicator Cluster Survey and 2018-2019 North Macedonia Roma Settlements Multiple Indicator Cluster Survey, Survey Findings Report.* Skopje, North Macedonia: State Statistical Office and UNICEF.

[6] Statistical Office of Montenegro (MONSTAT) and UNICEF. 2019. *2018 Montenegro Multiple Indicator Cluster Survey and 2018 Montenegro Roma Settlements Multiple Indicator Cluster Survey*, *Survey Findings Report*. Podgorica, Montenegro: MONSTAT and UNICEF.

explanatory dimensions because they reflect the resource-access logic central to the WEFE perspective. The Multidimensional Poverty Index (MPI) was used as the poverty measurement framework and provided the basis for constructing the binary dependent variable, through which households were classified as multidimensionally poor or non-poor. Logistic regression and machine-learning models were then applied to examine how WEFE-related variables, together with socio-economic predictors, are associated with that poverty outcome. The WEFE-related variables were not assigned separate manual weights in the machine-learning models, nor was any ex-ante prioritization imposed during model estimation. Rather, the WEFE framework informed the substantive selection of predictors, while their relative importance was determined empirically through model estimation, comparative model performance, and feature-importance analysis.

This methodological structure provides the basis for the subsequent empirical analysis by linking the conceptual role of the WEFE Nexus, the poverty-classification role of the MPI, and the explanatory role of the regression and machine-learning models.

Results and Discussion

Table 1

The dimensions, indicators, deprivation cut-offs, and weights of the Multidimensional Poverty Index

Dimensions of poverty	Indicator	Deprived of…	Weight
Education	Years of Schooling	No household member has completed five years of schooling.	1/6
	Child School Attendance	Any school-aged child is not attending school up to class 8.	1/6
Health	Child Mortality	Any child has died in the family.	1/6
	Nutrition	Any child for whom there is nutritional information is malnourished.	1/6
Living Standard (LS)	Electricity	The household has no electricity.	1/18
	Improved Sanitation	The household's sanitation facility is not improved (according to MDG guidelines) or improved but shared with other households.	1/18
	Improved Drinking Water	The household does not have access to improved drinking water (according to MDG guidelines), or safe drinking water is more than a 30-minute walk from home, roundtrip.	1/18
	Flooring	The household has a dirt, sand, or dung floor.	1/18
	Cooking Fuel	The household cooks with dung, wood, or charcoal.	1/18
	Assets ownership	The household does not own more than one telephone, mobile phone, radio, TV, bicycle, motorcycle/ scooter, car/ truck or animal driven cart or refrigerator, livestock (no cattle, no horse, less than two goats, or less than ten chicken).	1/18

Table 2

Multidimensional Poverty Index (MPI) and its Indices across Roma Settlements in Serbia, Roma Settlements, and Montenegro

	MPI			Poor Population (%)			Intensity (%)		
Roma Settlements	Overall	Rural	Urban	Overall	Rural	Urban	Overall	Rural	Urban
Serbia	3.95	4.27	3.79	10.11	10.80	9.74	39.13	39.52	38.91
North Macedonia	2.39	1.54	2.44	6.25	3.71	6.40	38.28	41.43	38.17
Montenegro	6.26	1.79	7.12	13.51	5.36	15.07	46.34	33.33	47.22

Table 2 Described the MPI and poverty indices in Roma settlements across Serbia, North Macedonia, and Montenegro, where Montenegro is facing the highest MPI not only overall but also among urban and rural areas. Overall, MPI and its indices are higher across Roma communities, indicating poverty, including

limited access to education, employment, housing, and other resources (Matache & Barbu, 2021; Matlovičová et al., 2022). This provides evidence of the need for a based approach to addressing structural inequities in resource access (Rasul, 2023). The data suggest that insufficient access to essential resources in Roma settlements can result in intense poverty, mostly in urban Montenegro. As with the goals of the current study, this baseline on Roma settlements highlights the potential of WEFE Nexus-compatible strategies to reduce poverty in marginalized communities. This research seeks to explore the mechanisms by which sustainable WEFE resource management can help bring about social inclusion and sustainable development for vulnerable groups in the Western Balkans by examining the interactions between resource access and poverty dimensions (Rochovská & Rusnáková, 2018; UNECE, 2022; Lekobane et al., 2024).

Please see Appendix Figure 2. for details regarding the percentage contribution of indicators to MPI across Roma settlements in Serbia, N. Macedonia, and Montenegro.

Table 3

Logistic Regression Analysis across Roma Settlements in Serbia, North Macedonia, and Montenegro

	Serbia (Roma Settlements)			N. Macedonia (Roma Settlements)			Montenegro (Roma Settlements)		
Variables	**Odds**	**SE**	**p-value**	**Odds**	**SE**	**p-value**	**Odds**	**SE**	**p-value**
Electric Lighting Access									
No	Reference								
Yes	0.059	0.016	0***	0.016	0.006	0***	0.1	0.041	0***
Sanitation Facility Emptied									
No	Reference								
Yes	1.704	0.216	0***	0.455	0.231	0.12	0.158	0.098	0.003***
Improved Water Access									
No	Reference								
Yes	0.788	0.242	0.437	-	-	-	0.067	0.048	0***
Reliable Water Supply									
No	Reference								
Yes	0.889	0.115	0.364	2.036	0.428	0.001***	1.549	0.6	0.259
Cooking Fuel Deprivation									
No	Reference								
Yes	6.105	1.165	0***	13.768	2.986	0***	4.892	1.266	0***
Male Household Head									
No	Reference								
Yes	0.613	0.095	0.002***	5.201	1.504	0***	1.971	0.642	0.037**
Number of Children Under Age 5	1.703	0.116	0***	0.76	0.084	0.013***	1.021	0.126	0.866
Age of Household Head	0.981	0.006	0.001***	0.95	0.007	0***	1.004	0.01	0.723
Ratio of Dependents	0.999	0.001	0.215	0.998	0.001	0.219	0.998	0.001	0.133
Has Bank Account									
No	Reference								
Yes	1.14	0.144	0.299	0.293	0.054	0***	1.052	0.234	0.818
Owns Dwelling									
No	Reference								
Yes	0.42	0.068	0***	0.644	0.132	0.031***	1.145	0.241	0.52
Education Level of Household Head									
Primary or None	Reference								
Secondary	0.39	0.085	0***	0.652	0.171	0.103	0.476	0.109	0.001***
Higher	.	.	.	.	.	.	1.129	0.566	0.809
Internet Access									
No	Reference								
Yes	0.125	0.017	0***	0.068	0.012	0***	0.106	0.03	0***
Urban Area									
No	Reference								
Yes	0.383	0.052	0***	0.569	0.187	0.086***	7.739	3.827	0***
Number of Rooms in Household	0.304	0.031	0***	0.409	0.057	0***	0.358	0.061	0***
Owns Agricultural Land									
No	Reference								
Yes	1.496	0.523	0.249	1	.	.	0.857	0.726	0.856
Owns Livestock or Animals									
No	Reference								
Yes	0.698	0.128	0.049**	2.322	0.66	0.003***	1.37	0.859	0.615
Region									

Belgrade	Reference			Vardar	Reference		South	Reference	
Vojvodina	1.62	0.338	0.021**	East (1.051)	0.904	0.953	Centre (.155)	0.079	0***
Sumadija and Western Serbia	0.842	0.202	0.475	South West (.380)	0.355	0.3	North (.183)	0.093	0.001***
Southern and Eastern Serbia	0.578	0.11	0.004***	South East (-)	-	-	Constant (37.411)	35.434	0***
Constant	78.163	43.59	0***	Pelagonija (3.123)	2.722	0.191			
				Polong (.306)	0.281	0.198			
				North East (2.336)	2.148	0.356			
				Skopje (2.519)	2.125	0.273			
				Constant (65.953)	65.959	0***			
Mean dependent var	0.077			0.066			0.117		
Pseudo r-squared	0.448			0.537			0.339		
Chi-square	1668.25			1310.387			385.467		
Akaike crit. (AIC)	2093.581			1171.675			793.264		
SD dependent var	0.266			0.248			0.321		
Number of obs	6876			5008			1579		
Prob > chi2	0			0			0		
Bayesian crit. (BIC)	2237.133			1315.088			905.919		

Note: Odds ratios are interpreted relative to the stated reference category for each variable.

Values above 1 indicate higher odds of multidimensional poverty relative to the reference group, whereas values below 1 indicate lower odds.

After constructing the MPI, we used binary logistic regression[7] to analyze the socio-economic factors influencing multidimensional poverty. Households with a deprivation score C_i equal to or above the cut-off k =33% were identified as multidimensionally poor, while those below the cut-off were classified as non-poor. This approach provides a useful framework for examining the correlates of multidimensional poverty across Roma settlements (Mohammed & Ab-Rahim, 2021; Ogwang, 2022; Abdiwahab, Menza, & Mohamed, 2024).

The logistic regression analysis for Roma settlements in Serbia, North Macedonia, and Montenegro reveals important socio-economic and resource-based correlates of multidimensional poverty within the WEFE Nexus framework. In particular, the analysis highlights the relevance of key resource-related variables, including electric lighting access, sanitation facility emptied, improved water access, reliable water supply, and cooking fuel deprivation, in relation to SDG 1, SDG 6, and SDG 7.

Electric lighting access shows a strong negative association with multidimensional poverty across all three countries, with odds ratios below 1 in Serbia (odds = 0.059, $p < 0.001$), North Macedonia (odds = 0.016, $p < 0.001$), and Montenegro (odds = 0.100, $p < 0.001$). This indicates that households using electricity for lighting are less likely to be multidimensionally poor than those relying on other lighting sources. At the same time, this relationship may also reflect broader differences in service access, housing quality, and infrastructure security within Roma settlements (El-Meligi et al., 2024).

[7] Alkire, S., Foster, J. E., Seth, S., Santos, M. E., Roche, J. M., and Ballon, P. (2015). Multidimensional Poverty Measurement and Analysis. Oxford: Oxford University Press, ch. 10.

$F(X'\beta) = \frac{e^{X'\beta}}{1+e^{X'\beta}} = \frac{\exp(X'\beta)}{1+\exp(X'\beta)}$, $C_i \geq k$ MPI Poor, $C_i < k$ MPI Non-Poor. Multidimensional Poor Household $= \begin{bmatrix} 0 & If\ No \\ 1 & If\ Yes \end{bmatrix}$

The odd ratio or relative risk is $\frac{P}{(1-p)}$ are measures the probability that y=1 relative to the probability that y=0

$$P = \frac{\exp(X'\beta)}{1+\exp(X'\beta)}, \quad \frac{P}{(1-p)} = \exp(X'\beta), \quad \ln\frac{P}{(1-p)} = X'\beta$$

The sanitation-related results also require cautious interpretation and differ across countries. In Serbia, sanitation facility emptied is positively associated with multidimensional poverty (odds = 1.704, $p < 0.001$), whereas in Montenegro the association is negative and statistically significant (odds = 0.158, $p = 0.003$). By contrast, the coefficient for North Macedonia is not statistically significant and therefore should not be interpreted as a robust effect. These differences suggest that the sanitation variable may reflect different underlying conditions across country contexts, including variation in infrastructure quality, maintenance practices, settlement form, and service provision, rather than a single uniform sanitation-poverty relationship (Matlovičová et al., 2022; UNECE, 2022).

Improved water access is also an important factor, particularly in Montenegro, where the odds ratio is strongly below 1 (odds = 0.067, $p < 0.001$), indicating a negative association with multidimensional poverty. This suggests that households with improved water access are less likely to experience overlapping deprivations, although the observed association may also reflect broader advantages in infrastructure provision, service reliability, and settlement conditions rather than the isolated effect of water access alone (Lange, 2018; Cheng, 2019; Zisopoulou & Panagoulia, 2021).

Reliable water supply displays a more context-specific pattern. In North Macedonia, the odds ratio is above 1 (odds = 2.036, $p < 0.001$), while the corresponding coefficients in Serbia and Montenegro are not statistically significant. This suggests that the relationship between reliable water supply and multidimensional poverty may be shaped by country-specific infrastructure arrangements, settlement characteristics, and service conditions rather than following a uniform pattern across the three study settings (Rodríguez-Pose & Stermšek, 2015).

The odds ratios for cooking fuel deprivation require particular attention. In Serbia (odds = 6.105, $p < 0.001$), North Macedonia (odds = 13.768, $p < 0.001$), and Montenegro (odds = 4.892, $p < 0.001$), the results indicate that households experiencing cooking fuel deprivation have substantially higher odds of multidimensional poverty relative to the reference category. This association likely captures broader forms of structural disadvantage, including limited access to clean energy infrastructure, lower housing quality, financial constraints, and settlement-level exclusion, rather than the isolated effect of cooking fuel alone (Bouzarovski, 2017; Rasul, 2023).

Several socio-economic variables also emerge as important correlates of multidimensional poverty. In North Macedonia, having a bank account is negatively associated with poverty (odds = 0.293, $p < 0.001$), suggesting that households with access to formal financial services are less likely to be multidimensionally poor. Similarly, internet access is negatively associated with poverty across all three countries, reinforcing the link between connectivity, access to opportunities, and social inclusion in marginalized communities (Choudrie et al., 2018; Yakubi et al., 2022; Kovač et al., 2024). The number of rooms in the household is also consistently negatively associated with poverty, suggesting that better housing conditions are linked to improved living standards (Vuksanović-Macura, 2012; Vuksanović-Macura & Miščević, 2021).

Other household-level factors show more context-dependent effects. For example, a male household head is negatively associated with multidimensional poverty in Serbia (odds = 0.613, $p = 0.002$), but positively associated with poverty in North Macedonia (odds = 5.201, $p < 0.001$), while in Montenegro the relationship is weaker though still positive (odds = 1.971, $p = 0.037$). Similarly, urban area is negatively associated with poverty in Serbia (odds = 0.383, $p < 0.001$), but positively associated in Montenegro (odds = 7.739, $p < 0.001$), indicating that the social and spatial determinants of deprivation differ across the three country settings. These contrasts suggest that household structure, settlement patterns, and local opportunity structures may shape poverty in different ways across Roma communities.

Overall, the logistic regression findings should be interpreted as associational rather than causal. Several WEFE-related variables likely capture broader latent dimensions of deprivation, including infrastructure deficits, housing precarity, service exclusion, and chronic socio-economic disadvantage within Roma settlements. Accordingly, the odds ratios reported in Table 3 are best understood as indicators of how specific household conditions are statistically linked to multidimensional poverty, rather than as direct estimates of the isolated causal effect of a single factor.

Table 4

Comparative Performance of Machine Learning Models across Roma Settlements in Serbia, North Macedonia, and Montenegro

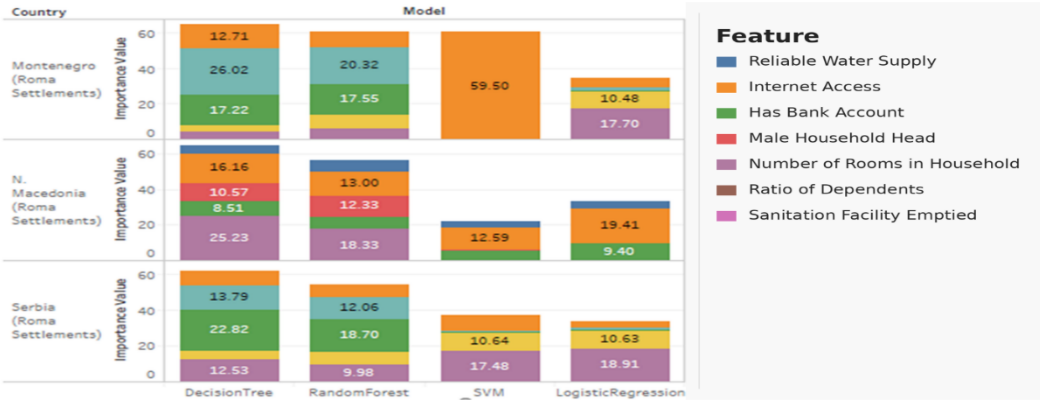
	Random Forest	Decision Tree	SVM	Logistic Regression	Gradient Boosting	Ada Boost	XG Boost	Light GBM	Extra Trees
Serbia (Roma Settlements)									
Accuracy	0.998547	0.998547	0.941134	0.946221	0.962209	0.946948	0.998547	0.998547	0.998547
Precision	0.979798	0.979798	0.863636	0.709091	0.894737	0.706897	0.979798	0.979798	0.979798
Recall	1	1	0.195876	0.402062	0.525773	0.42268	1	1	1
F1-Score	0.989796	0.989796	0.319328	0.513158	0.662338	0.529032	0.989796	0.989796	0.989796
ROC AUC	0.999976	0.999992	0.92133	0.927239	0.969556	0.949203	0.999976	0.999992	0.999992
North Macedonia (Roma Settlements)									
Accuracy	0.995079	0.996063	0.951772	0.951772	0.969488	0.950787	0.995079	0.995079	0.996063
Precision	1	1	0.823529	0.761905	0.97619	0.733333	1	1	1
Recall	0.929577	0.943662	0.394366	0.450704	0.577465	0.464789	0.929577	0.929577	0.943662
F1-Score	0.963504	0.971014	0.533333	0.566372	0.725664	0.568966	0.963504	0.963504	0.971014
ROC AUC	0.998174	0.971831	0.950876	0.960176	0.983732	0.967598	0.991743	0.991967	0.985804
Montenegro (Roma Settlements)									
Accuracy	0.996835	0.981013	0.889241	0.908228	0.974684	0.889241	0.981013	0.981013	0.996835
Precision	1	0.871795	0.5	0.75	1	0.5	0.871795	0.871795	1
Recall	0.971429	0.971429	0.171429	0.257143	0.771429	0.257143	0.971429	0.971429	0.971429
F1-Score	0.985507	0.918919	0.255319	0.382979	0.870968	0.339623	0.918919	0.918919	0.985507
ROC AUC	0.999288	0.976817	0.877885	0.879309	0.99878	0.953533	0.998983	0.998983	0.999492

Table 4 presents the comparative predictive performance of a broad set of machine-learning algorithms estimated for each country context. Overall, ensemble-based models, particularly Random Forest and Extra Trees, demonstrate the strongest and most consistent predictive performance across the three study settings. These models achieve very high levels of accuracy and ROC AUC, suggesting that they are well suited to capturing the complex socio-economic and resource-related patterns associated with multidimensional poverty in Roma settlements (Alkire & Foster, 2011; Cheng, 2019; Hassan, 2024).

The machine-learning results complement the logistic regression findings by showing that multidimensional poverty is shaped by multiple interacting predictors rather than by isolated socio-economic factors alone. In this regard, ensemble models are especially useful because they can capture non-linear relationships and interaction effects that may not be fully reflected in the logistic specification (Cheng, 2019; Kambuya, 2020). At the same time, the comparative model results indicate that performance patterns are not identical across countries, which points to meaningful contextual variation in poverty-related risks and supports the need for country-sensitive poverty analysis within the WEFE Nexus framework (Ozden & Guleryuz, 2022; UNECE, 2022).

Beyond predictive performance, the machine-learning analysis also provides insight into which socio-economic and WEFE-related variables contribute most strongly to multidimensional poverty classification. Figure 1 therefore summarizes the most influential predictors across Serbia, North Macedonia, and Montenegro and helps clarify how the relative prominence of these variables differs across country contexts. Figure 1. Feature importance across machine-learning models for multidimensional poverty prediction in Roma settlements in Serbia, North Macedonia, and Montenegro. The figure presents the relative contribution of the most influential socio-economic and WEFE-related predictors across the classification models. Higher values indicate greater contribution of a given predictor to model-based poverty classification within each country setting.

Figure 1 provides a clearer interpretation of the machine-learning results by showing which socio-economic and WEFE-related variables contribute most strongly to poverty prediction across the three country settings. In Montenegro, internet access emerges as the most influential predictor, indicating that digital exclusion is closely associated with multidimensional poverty in Roma settlements. This finding reinforces earlier evidence that digital connectivity is increasingly tied to access to information, services, and livelihood opportunities in marginalized communities (Sarma & Pais, 2011; World Bank Group, 2016; Yakubi et al., 2022). Water- and sanitation-related variables also remain prominent, suggesting that deficits in basic service access continue to shape deprivation in important ways, consistent with the WEFE Nexus emphasis on



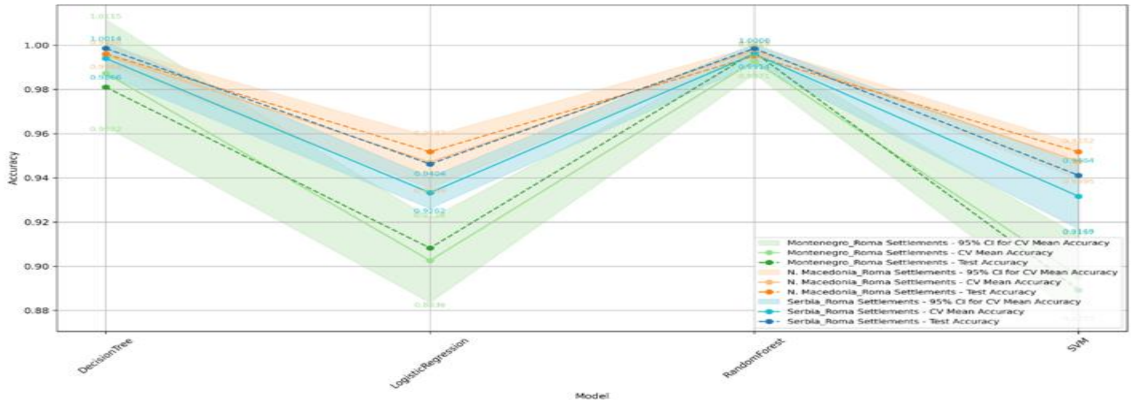
resource-linked poverty pathways and with SDG 6 priorities (Wang, 2017; Cheng, 2019; UNECE, 2022).

In North Macedonia, the number of rooms in the household and the ratio of dependents appear among the most important predictors, highlighting the role of household composition, overcrowding, and housing conditions in explaining poverty vulnerability. These patterns are consistent with the broader literature showing that constrained living space and higher dependency burdens intensify multidimensional deprivation in marginalized communities (Huang, 2017). Internet access and sanitation facility emptied also retain considerable importance, suggesting that material infrastructure and social connectivity jointly influence multidimensional deprivation in this setting (El-Meligi et al., 2024).

In Serbia, the most influential predictors include the number of rooms in the household, sanitation facility emptied, has bank account, and male household head. These findings suggest that poverty in Roma settlements is shaped not only by deficiencies in basic living conditions, but also by broader socio-economic vulnerability and unequal access to financial resources. This is consistent with earlier work showing that household structure, gendered vulnerability, and financial access are important dimensions of poverty resilience and economic stability (Demirgüç-Kunt & Klapper, 2012; Ravallion, 2015; Robayo-Abril & Millan, 2019; Jasini, 2024).

Taken together, the feature-importance results show that multidimensional poverty across the Western Balkan Roma settlements is driven by a combination of shared structural disadvantages and country-specific patterns of deprivation. This strengthens the interpretive value of the machine-learning analysis and supports the broader WEFE Nexus perspective that resource inequalities, housing constraints, and socio-economic exclusion are mutually reinforcing dimensions of poverty (Cheng, 2019; UNECE, 2022; El-Meligi et al., 2024).

Figure 2



Model performance with test and cross-validation mean accuracies including confidence intervals across Serbia, North Macedonia, and Montenegro. Random Forest and Extra Trees display the highest and most stable predictive performance across the three country settings. Their relatively tight confidence intervals indicate not only strong classification accuracy but also greater model stability. By contrast, simpler models such as SVM and Logistic Regression show comparatively lower performance and greater variability in some settings, especially in Serbia and North Macedonia. These findings further support the usefulness of ensemble models for capturing complex poverty-related patterns in underserved communities and for strengthening multidimensional poverty analysis in contexts characterized by non-linear socio-economic interactions (Ozden & Guleryuz, 2022; Usmanova et al., 2022; El-Meligi et al., 2024).

Logistic Regression and Machine Learning for Poverty Prediction within the WEFE Nexus

This paper integrates logistic regression and machine-learning models to examine how socio-economic and resource-based factors are associated with multidimensional poverty across Roma settlements in Serbia, North Macedonia, and Montenegro. Within this framework, the WEFE Nexus is used to interpret the role of interconnected household resource conditions, particularly electric lighting access, sanitation facility emptied, improved water access, reliable water supply, and cooking fuel deprivation, in shaping multidimensional poverty outcomes (Zisopoulou & Panagoulia, 2021). These dimensions are closely connected to SDG 1, SDG 6, SDG 7, and broader goals related to social sustainability.

In this design, the WEFE Nexus serves as the conceptual lens for selecting the relevant resource-related predictors, the MPI provides the poverty classification outcome, and the regression and machine-learning models identify how these dimensions are statistically linked at the household level.

Results from the logistic regression analysis show that several WEFE-related variables are strongly associated with multidimensional poverty. Electric lighting access is negatively associated with poverty across all three countries, indicating that households with electricity-based lighting are less likely to be multidimensionally poor. Improved water access in Montenegro is also negatively associated with poverty, while cooking fuel deprivation is positively associated with multidimensional poverty in Serbia, North Macedonia, and Montenegro. These patterns highlight the importance of household resource conditions in shaping deprivation, while also suggesting that broader contextual factors such as infrastructure quality, settlement form, service access, and country-specific institutional conditions may underline these relationships (Cheng, 2019; Matache & Barbu, 2021).

The machine-learning results complement the logistic findings by showing that ensemble methods such as Random Forest, Extra Trees, and XGBoost are highly effective in capturing the complex socio-economic and resource-related patterns associated with multidimensional poverty. The very high accuracy and ROC AUC values reported across countries indicate the usefulness of these models for poverty classification within the WEFE Nexus framework (Alkire & Foster, 2011). The feature-importance analysis further strengthens the interpretation by identifying the predictors that contribute most strongly to poverty classification in each country context, thereby supporting more targeted and context-sensitive policy discussion (Ahmadi et al., 2024).

Discussion

Additional insights on model performance and resource priorities are provided through the country-specific analysis. In Serbia, the strongest predictors include number of rooms in the household, sanitation facility emptied, has bank account, and male household head, indicating that both material living conditions and socio-economic vulnerability are important in shaping poverty outcomes. In North Macedonia, the results place greater emphasis on household composition and housing conditions, particularly number of rooms in the household and ratio of dependents, alongside internet access and sanitation-related conditions. In Montenegro, internet access emerges as a particularly important predictor, together with reliable water supply and other service-related conditions. These patterns suggest that although multidimensional poverty in Roma settlements reflects common structural disadvantages, the relative prominence of specific predictors differs across the three country settings (Cheng, 2019).

These cross-country differences are likely shaped by broader contextual factors rather than by a single common poverty mechanism. In Montenegro, the stronger role of internet access may reflect the growing importance of digital connectivity for access to services, information, and opportunities in smaller and more spatially uneven settlement contexts. In Serbia, the greater prominence of sanitation-related conditions and access to formal financial services may point to the continued importance of basic service deficits alongside economic exclusion in larger and more internally differentiated Roma communities. In North Macedonia, the stronger role of household composition and housing-related constraints suggests that overcrowding, dependency burden, and family-level resource pressures remain especially important in shaping poverty vulnerability. These contrasts indicate that multidimensional poverty is influenced not only by shared deprivation, but also by differences in policy environments, infrastructure provision, historical marginalization, and local socio-economic conditions across the three countries (Rochovská & Rusnáková, 2018; Orton et al., 2019; World Bank, 2019).

The findings of the study emphasize the need for targeted policies that respond to differences in the socio-economic and resource-based challenges faced by Roma settlements across the Western Balkans. At the local-government level, the results suggest that interventions should prioritize household service deficits that are most strongly associated with multidimensional poverty in each country's context. In areas where electric lighting access and cooking fuel deprivation are especially important, municipal and national actors should expand access to cleaner household energy through subsidized electricity connections, safer cooking technologies, and settlement-level upgrading of basic infrastructure. In locations where improved water access, reliable water supply, and sanitation-related conditions remain prominent, local governments and service providers should prioritize piped service extension, regularized sanitation maintenance, and community-level investment in safe waste and wastewater systems (Bhabha et al., 2017; World Bank, 2019).

The findings also point to a practical role for NGOs and community-based organizations. Where internet access and having a bank account emerge as strong predictors, NGOs can support digital inclusion through community access points, digital literacy initiatives, and assistance with access to online public services, while also facilitating financial inclusion through documentation support, account-opening assistance, and outreach to underserved households. In addition, because housing-related variables such as number of rooms and household crowding remain important across settings, poverty-reduction strategies should include housing improvement, overcrowding reduction, and dwelling regularization measures as part of a broader multidimensional intervention framework (Alkire & Foster, 2011; Matache & Barbu, 2021; Lekobane et al., 2024; Ahmadi et al., 2024).

More broadly, the machine-learning results can help policymakers and NGOs prioritize interventions according to the strongest poverty-related predictors in each country setting rather than applying a uniform strategy across all Roma settlements. In this sense, the study supports a more operational use of poverty analytics: Serbia may require stronger emphasis on sanitation-related deficits and financial exclusion, North Macedonia on housing pressure and household dependency burden, and Montenegro on internet connectivity and service-related access conditions. Such a targeted approach would improve the practical value of resource allocation by aligning interventions more closely with the country-specific deprivation structures identified in the analysis.

Conclusion

The consistent identification of key WEFE-related and socio-economic factors across Roma settlements in Serbia, North Macedonia, and Montenegro indicates the need for an integrated approach to multidimensional poverty reduction. The findings suggest that household resource conditions, housing quality, service access, and socio-economic inclusion are closely linked to poverty outcomes in these communities. In this respect, the study supports the broader WEFE Nexus perspective by showing that multidimensional poverty is shaped by interconnected forms of deprivation rather than by isolated deficits alone.

By combining logistic regression with machine-learning analysis, the study provides a useful framework for identifying both shared and country-specific predictors of multidimensional poverty in Roma settlements. These results offer a stronger basis for SDG-aligned policy thinking in the Western Balkans, especially in relation to resource access, infrastructure improvement, housing conditions, and financial and digital inclusion. Rather than treating poverty as a single-dimensional problem, the study highlights the value of integrated and context-sensitive approaches to social sustainability and poverty reduction in marginalized communities, while also underscoring that effective interventions must be responsive to differences in regional policy environments, infrastructure conditions, and historical patterns of exclusion across the three countries.

The findings also suggest that local governments and NGOs can use country-specific predictor patterns identified through the machine-learning analysis to guide more practical and targeted interventions in energy, sanitation, housing, financial inclusion, and digital access.

Acknowledgement:
This research is part of the 101136834 – CROSS-REIS – HORIZON-WIDERA-2023-ACCESS-03 project, funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the European Research Executive Agency can be held responsible for them.

References

Abdiwahab, B., Menza, M., & Mohamed, A. A. (2024). Multidimensional Poverty and its Determinants in Somalia: A Household Level Analysis. *International Journal of Developing Country Studies*, *6*(1), 36-57.

Ahmadi, M., Mosazadeh, N., & Hatami, H. (2024). Identifying the causes of women's tendency to go to koolbari in border villages: Case study of border villages in Kermanshah province. *Journal of Rural Research*. https://jrur.ut.ac.ir/article_99323_en.html?lang=en

Alkire, S., & Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics, 95*(7 8), 476-487. https://doi.org/10.1016/j.jpubeco.2010.11.006

Alkire, S., & Santos, M. E. (2013). A multidimensional approach: Poverty measurement & beyond. *Social indicators research*, *112*(2), 239-257. https:// doi.org/10.1007/s11205-013-0257-3

Al-Saidi, M., & Elagib, N. A. (2017). Towards understanding the integrative approach of the water, energy, and food nexus. *Science of the Total Environment, 574*, 1131-1139. https://doi.org/10.1016/j.scitotenv.2016.09.046

Bazilian, M., Rogner, H., Howells, M., Hermann, S., Arent, D., Gielen, D., ... & Yumkella, K. K. (2011). Considering the energy, water, and food nexus: Towards an integrated modeling approach. *Energy Policy, 39*(12), 7896 7906. https://doi.org/10.1016/j.enpol.2011.09.039

Bhabha, J., Fuller, A., Matache, M., Vranješević, J., Chernoff, M. C., Spasić, B., & Ivanis, J. (2017). Reclaiming adolescence: A Roma youth perspective. *Harvard Educational Review, 87*(2), 186-224. https://doi.org/10.17763/1943-5045-87.2.186

Bieber, F. (2005). *Post-war Bosnia: Ethnicity, inequality and public sector governance*. Springer. https://doi.org/10.1057/9780230501379

Bojović, M., Mrkonjić, Z., & Vukelić, I. (2024). Agroforestry systems and forest resources as a potential for sustainable energy development in the western Balkan region. *Energy, Sustainability, and Society, 14*(68). https://doi.org/10.1186/s13705-024-00502-y

Bouzarovski, S., & Tirado Herrero, S. (2017). The energy divide: Integrating energy transitions, regional inequalities and poverty trends in the European Union. *European Urban and Regional Studies, 24*(1), 69-86. https://doi.org/10.1177/0969776415596449

Cheng, Xin (2019). Modelling environment and poverty for sustainable development. Macquarie University. Thesis. https://doi.org/10.25949/27086674.v1

Choudrie, J., Tsatsou, P., & Kurnia, S. (Eds.). (2018). Social inclusion and usability of ICT-enabled services. New York City: Routledge. https://doi.org/10.4324/9781315677316

Demirgüç-Kunt, A., & Klapper, L. F. (2012). Financial inclusion in Africa: An overview (World Bank Policy Research Working Paper No. 6088). World Bank. Available at SSRN: https://ssrn.com/abstract=2084599

Do, M. H. (2024). Rural development policies in Southeast Asia: Impacts on household livelihoods, welfare, and rural transformation. *University of Hannover Repository*. Retrieved from https://www.repo.uni-hannover.de/handle/123456789/18135

El-Meligi, A., Ferreira, V., Nechifor, V., & Ferrari, E. (2024). Application of environmentally extended social accounting matrix-based indicators to address socio-economic and environmental targets in Cameroon. *SSRN Electronic Journal*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4987914

Hassan, A. A., Muse, A. H., & Chesneau, C. (2024). Machine learning study using 2020 SDHS data to determine poverty determinants in Somalia. *Scientific Reports, 14*, 5956. https://doi.org/10.1038/s41598-024-56466-8

Huang, S., Feng, Q., Lu, Z., Wen, X., & Deo, R. C. (2017). Trend analysis of water poverty index for assessment of water stress and water management policies: A case study in the Hexi Corridor, China. *Sustainability, 9*(5), 756. https://doi.org/10.3390/su9050756

Ivanov, A., Keller, S., and Till-Tentschert, U. (2015). "Roma Poverty and Deprivation: The Need for Multidimensional Anti-Poverty Measures." OPHI Working Papers 96, University of Oxford.

Jasini, K. (2024). The regime of inequality affecting the Roma community in Albania. *Academicus International Scientific Journal*, *15*(29), 88-99.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science, 353*(6301), 790-794. https://doi.org/10.1126/science.aaf7894

Kambuya, P. (2020). Better model selection for poverty targeting through machine learning: A case study in Thailand. *Thailand and The World Economy*, *38*(1), 91-116.

Kovač, N., Żmija, K., Roy, J. K., Kusa, R., & Duda, J. (2024). Digital divide and digitalization in Europe: A bibliometric analysis. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 19*(2), 463–520. https://doi.org/10.24136/eq.2899

Lange, M. A. (2018). The Impacts of Climate Change in the MENA Region and the Water-Energy Nexus. http://dx.doi.org/10.20944/preprints201810.0197.v1

Lekobane, K. R., & Samboma, T. A. (2024). Children with disabilities: Left behind or forgotten? Empirical evidence from Botswana using multidimensional poverty measures. *Children and Youth Services Review*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0190740924005577

Matache, M., & Barbu, S. (2021). Assessing racialized poverty: The case of Romani people in the European Union. In *Research handbook on human Rights and poverty* (pp. 192-210). Edward Elgar Publishing. https://doi.org/10.4337/9781788977517.00023

Matlovičová, K., Kolesárová, J., Demková, M., Kostilníková, K., Mocák, P., Pachura, P., & Payne, M. (2022). Stimulating poverty alleviation by developing tourism in marginalised Roma communities: A case study of the Central Spiš Region (Slovakia). *Land*, *11*(10), 1689. https://doi.org/10.3390/land11101689

Mikula, N., Lugeri, F., & Leip, A. (2018). The water-energy-food nexus and multidimensional poverty: A review. *Sustainable Development Journal, 26*(6), 736-748. https://doi.org/10.1002/sd.1872

Mohammed, M., & Ab-Rahim, R. (2021). Determinants of Multidimensional Poverty Index of Niger State, Nigeria. *International Journal of Academic Research in Business and Social Sciences*, *11*(14), 95–108. http://dx.doi.org/10.6007/IJARBSS/v9-i7/6137

Morar, B., Gresham, D., Angelicheva, D., Tournev, I., Gooding, R., Guergueltcheva, V., Schmidt, C., Abicht, A., Lochmüller, H., Tordai, A., Kalmár, L., Nagy, M., Karcagi, V., Jeanpierre, M., Herczegfalvi, A., Beeson, D.,Venkataraman, V., Carter, K. W., Reeve, J., de Pablo, R., Kucˇinskas, V., & Kalaydjieva, L. (2004). Mutation history of the Roma/Gypsies. *The American Journal of Human Genetics, 75*(4), 596-609.

Nyhan, K., Harsono, D., Atre, S., Peterson, H., & Nyhan, K. (2024). A scoping review of factors associated with HIV acquisition in the context of humanitarian crises. *AIDS and Behavior*. Retrieved from https://link.springer.com/article/10.1007/s10461-024-04504-x

Ogwang, T. (2022). The regression approach to the measurement and decomposition of the multidimensional Watts poverty index. *J Econ Inequal* 20, 951–973. https://doi.org/10.1007/s10888-022-09531-z

Orton, L., Anderson de Cuevas, R., Stojanovski, K., Gamella, J. F., Greenfields, M., La Parra, D., Marcu, O., Matras, Y., Donert, C., Frost, D., Robinson, J., Rosenhaft, E., Salway, S., Sheard, S., Such, E., Taylor-Robinson, D., & Whitehead, M. (2019). Roma populations and health inequalities: A new perspective. *International Journal of Human Rights in Healthcare, 12*(5), 319–327. https://doi.org/10.1108/IJHRH-01-2019-0004

Ozden, E., & Guleryuz, D. (2022). Optimized machine learning algorithms for investigating the relationship between economic development and human capital. *Computational Economics*, *60*(1), 347-373. https://doi.org/10.1007/s10614-021-10194-7

Rashid, N. K. A., & Kulub, N. (2026). The Influence of Multidimensional Poverty Towards Life Satisfaction. *FWU Journal of Social Sciences*, *20*(1), 1.

Rasul, G. (2016). Managing the food, water, and energy nexus for achieving the Sustainable Development Goals in South Asia. *Environmental Development, 18*, 14-25. https://doi.org/10.1016/j.envdev.2015.12.001

Rasul, G. (2023). Poverty, environment, and development nexus: Addressing resource-based poverty in South Asia. *Sustainable Development Journal, 33*(1), 22-36. https://doi.org/10.1002/sd.2339

Ravallion, M. (2015). The economics of poverty: History, measurement, and policy. Oxford University Press.

Robayo-Abril, M., & Millan, N. (2019). Breaking the cycle of Roma exclusion in the Western Balkans. World Bank. http://hdl.handle.net/10986/31393

Ringler, C., Bhaduri, A., & Lawford, R. (2013). The nexus across water, energy, land and food (WELF): potential for improved resource use efficiency?. *Current Opinion in Environmental Sustainability*, *5*(6), 617-624. http://dx.doi.org/10.1016/j.cosust.2013.11.002

Rochovská, A., & Rusnáková, J. (2018). Poverty, segregation, and social exclusion of Roma communities in Slovakia. *Bulletin of GeoFigurey. Socio-economic Series, 42*(42), 195–212. https://doi.org/10.2478/bog-2018-0039

Rodríguez-Pose, A., & Stermšek, M. (2015). The Economics of Secession: Analysing the Economic Impact of the Collapse of the Former Yugoslavia. *Territory, Politics, Governance, 3*(1), 73–96. https://doi.org/10.1080/21622671.2014.965729

Salas-Rojo, P., Rodríguez, J.G. (2022). Inheritances and wealth inequality: a machine learning approach. *J Econ Inequal* 20, 27–51 https://doi.org/10.1007/s10888-022-09528-8

Sarma, M., & Pais, J. (2011). Financial inclusion and development. *Journal of International Development, 23*(5), 613-628. https://doi.org/10.1002/jid.1698

Trbojević, M., Jovanović, D., & Đurđević, D. (2024). Security policies and sustainable development in the Western Balkan region beyond 2022: current status, challenges, and prospects. *Energy, Sustainability and Society*, *14*(1), 55. https://doi.org/10.1186/s13705-024-00486-9

United Nations Economic Commission for Europe (UNECE). (2022). Resource scarcity and its impacts on poverty in the Western Balkans: Challenges and solutions. *United Nations Economic Commission for Europe*.

Usmanova, A., Aziz, A., Rakhmonov, D., & Osamy, W. (2022). Utilities of artificial intelligence in poverty prediction: A review. *Sustainability, 14*(21), 14238. https://doi.org/10.3390/su142114238

Vuksanović-Macura, Z. (2012). The mapping and enumeration of informal Roma settlements in Serbia. *Environment and Urbanization, 24*(2), 685-705. https://doi.org/10.1177/0956247812451809

Vuksanović-Macura, Z., & Miščević, I. (2021). Excluded communities and participatory land-use planning: Experience from informal Roma settlements in Serbia. *Environment and Urbanization, 33*(2), 456-477. https://doi.org/10.1177/09562478211024095

Wang, J., Li, Y., Huang, J., Yan, T., & Sun, T. (2017). Growing water scarcity, food security and government responses in China. *Global Food Security, 14*, 9-17. https://doi.org/10.1016/j.gfs.2017.01.003

World Bank. (2019). The problem is they see us as a whole: Explaining gender and ethnicity-based inequalities among the Roma in Serbia. World Bank.

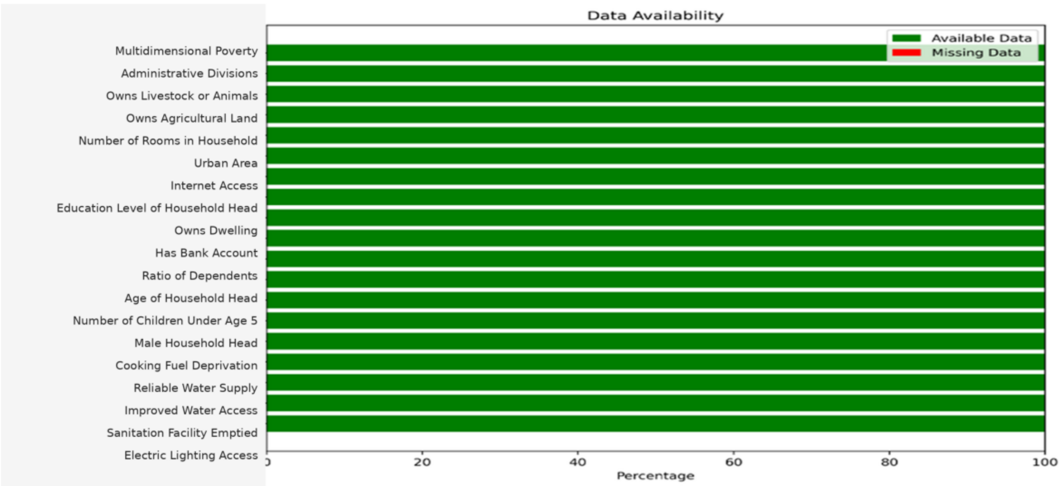
World Bank Group. (2016). World development report 2016: Digital dividends. World Bank Publications.

Yakubi, Y. A. Y., Basuki, B., Purwono, R., & Usman, I. (2022). The impact of digital technology and business regulations on financial inclusion and socio-economic development in low-income countries. *SAGE Open, 12*(3). https://doi.org/10.1177/21582440221116112

Zisopoulou, K., & Panagoulia, D. (2021). An in-depth analysis of physical blue and green water scarcity in agriculture in terms of causes and events and perceived amenability to economic interpretation. *Water, 13*(12), 1693. https://doi.org/10.3390/w13121693

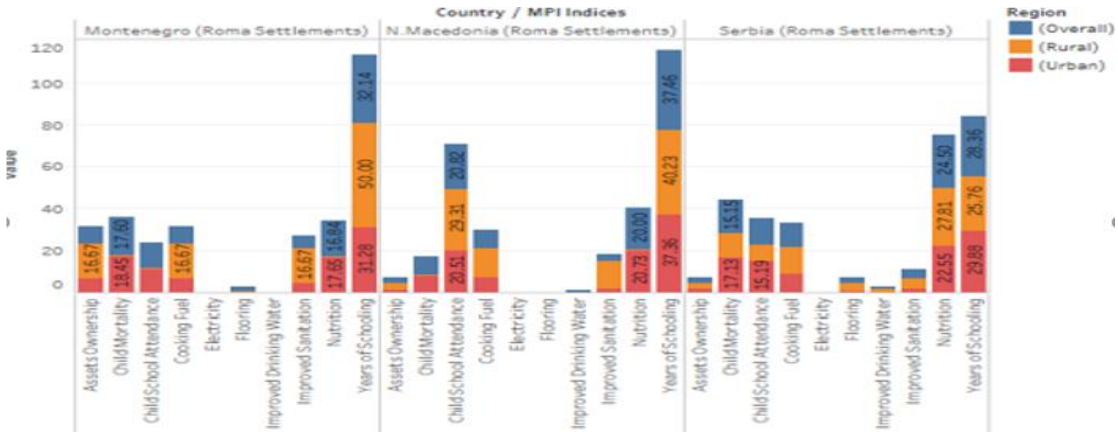
Appendix

Appendix Figure 1. Data Availability



Appendix Figure 1 illustrates the availability of data for the selected socio-economic and resource-related variables under study, where complete data (100% availability) exists for each indicator. The high level of data availability makes it possible to undertake multidimensional poverty analysis and resource access, which in turn supports the development of robust insights into the contribution to poverty dynamics of these domains (water supply, sanitation, energy) within the WEFE Nexus framework.

Figure 2
Percentage Contribution of indicators to MPI across Roma Settlements in Serbia, North Macedonia, and Montenegro
The



Appendix Figure 2. Shows how different indicators contribute differently to the Multidimensional Poverty Index (MPI) at Roma settlements in Serbia, North Macedonia, and Montenegro by indicating the overall, rural and urban regions.

Child school attendance, years of schooling, and access to improved sanitation all play important roles in MPI values in all three countries, consistent with the point that education and access to sanitation play a crucial role in multidimensional poverty (Alkire and Foster, 2011). The almost 50 percent contribution of school attendance in rural Montenegro clearly shows educational deprivation. A similar trend is observed in North Macedonia as school attendance and years of schooling contribute similarly, but especially in rural areas, providing evidence of the role of education in poverty reduction (UNECE, 2022). The results highlight the emphasis on targeted interventions in education and sanitation within Roma settlements to diminish poverty across the regions (Ferrer, & Garcia, 2023).

Appendix Figure 3. ROC Figures from both traditional and machine learning models

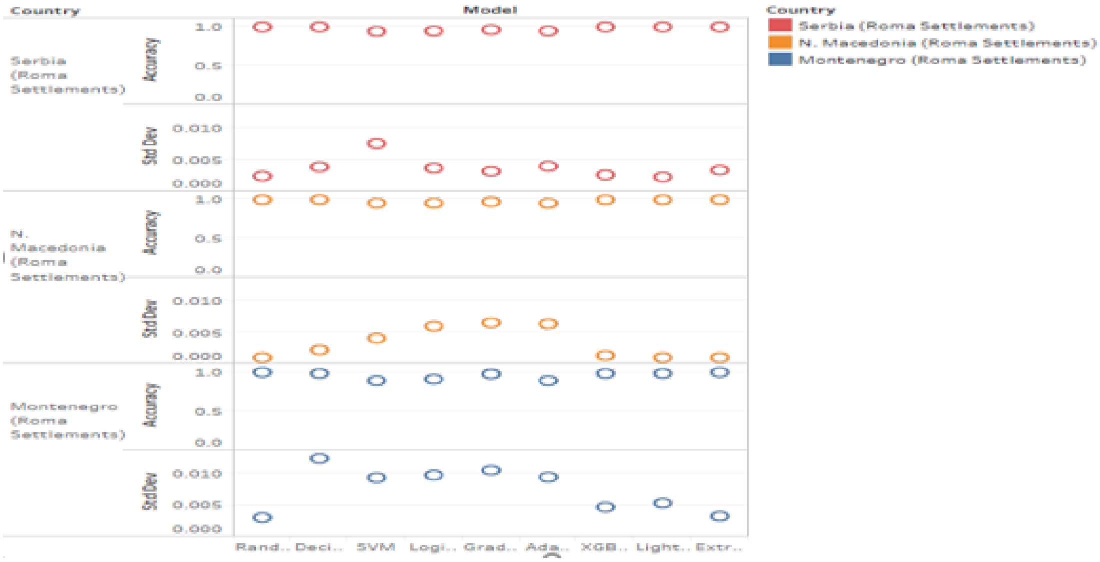
Comparative analysis of AUC ROC curves shows pronounced differences in the accuracy of model prediction of multidimensional poverty between logistic regression and machine learning models in Roma settlements in Serbia, North Macedonia, and Montenegro. The AUC ROC curves for the logistic regression were strong (0.93) in Serbia classifying between poverty and non-poverty cases. Although in machine learning these models have near-perfect predictive accuracy with an AUC of around 1.0, this is indicative of their robustness to learning complex, non-linear socio-economic relationships. Logistic regression reaches an AUC of about 0.88 for Montenegro, a moderate value if logistic regression oversimplifies poverty determinants too much, while machine learning models like Gradient Boosting, Random Forest, and Light GBM achieve near-perfect AUCs. Across the three countries, overall, machine learning models show consistently stronger AUC scores than logistic regression, indicating that they can better identify poverty determinants according to the WEFE Nexus framework. This performance advantage suggests that machine learning can be used to inform directed, resources-based poverty alleviation strategies that are aligned with social sustainability and SDG goals.

Tanble **1.** *Descriptive Statistics for Roma Settlements in Serbia, North Macedonia, and Montenegro*

Variables	Serbia (Roma Settlements) Obs.	Mean	S D	Min	Max	N. Macedonia (Roma Settlements) Obs.	Mean	S D	Min	Max	Montenegro (Roma Settlements) Obs.	Mean	S D	Min	Max
Reliable Lighting	6876	.98	.13	0	1	5078	.98	.11	0	1	1579	.96	.17	0	1
Sanitation Maintenance	6876	.3	.45	0	1	5078	.01	.12	0	1	1579	.06	.24	0	1
Safe Water Access	6876	.98	.13	0	1	5078	.99	.03	0	1	1579	.98	.11	0	1
Consistent Water Supply	6876	.70	.45	0	1	5078	.70	.45	0	1	1579	.90	.29	0	1
Energy for Cooking	6876	.56	.49	0	1	5078	.36	.48	0	1	1579	.50	.5	0	1
Gender of Household Head	6876	.83	.37	0	1	5078	.83	.36	0	1	1579	.86	.33	0	1
Number of Children Under Age 5	6876	.89	.97	0	4	5078	.72	.91	0	5	1579	1.28	1.21	0	6
Age of Household Head	6876	46.27	13.11	17	95	5078	46.35	12.6	18	82	1579	40.43	12.23	15	95
Ratio of Dependents	6876	84.98	84.18	0	800	5078	73.66	65.60	0	500	1579	131.95	110.45	0	700
Financial Inclusion	6876	.62	.48	0	1	5078	.87	.33	0	1	1579	.46	.49	0	1
Ownership Status of Household	6876	.89	.31	0	1	5078	.88	.31	0	1	1579	.63	.48	0	1
Education Level of HH Head	6876	.20	.40	0	1	5078	.19	.39	0	1	1579	.71	.57	0	2
Digital Access	6876	.75	.43	0	1	5078	.79	.40	0	1	1579	.51	.5	0	1
Area	6876	.67	.46	0	1	5078	.93	.24	0	1	1579	.80	.39	0	1
Number of Rooms in Household	6876	2.25	.97	1	8	5078	2.11	.89	1	7	1579	2.12	1.17	1	8
Ownership of Agricultural Land	6876	.03	.17	0	1	5078	.01	.11	0	1	1579	.01	.13	0	1
Ownership of Livestock or Animals	6876	.10	.30	0	1	5078	.03	.18	0	1	1579	.03	.17	0	1
Region	6876	1.99	1.20	0	3	5078	5.06	2.26	0	7	1579	1.00	.46	0	2

Appendix Figure 4. Cross-Validation Accuracy with Standard Deviation across Serbia, North Macedonia, and Montenegro

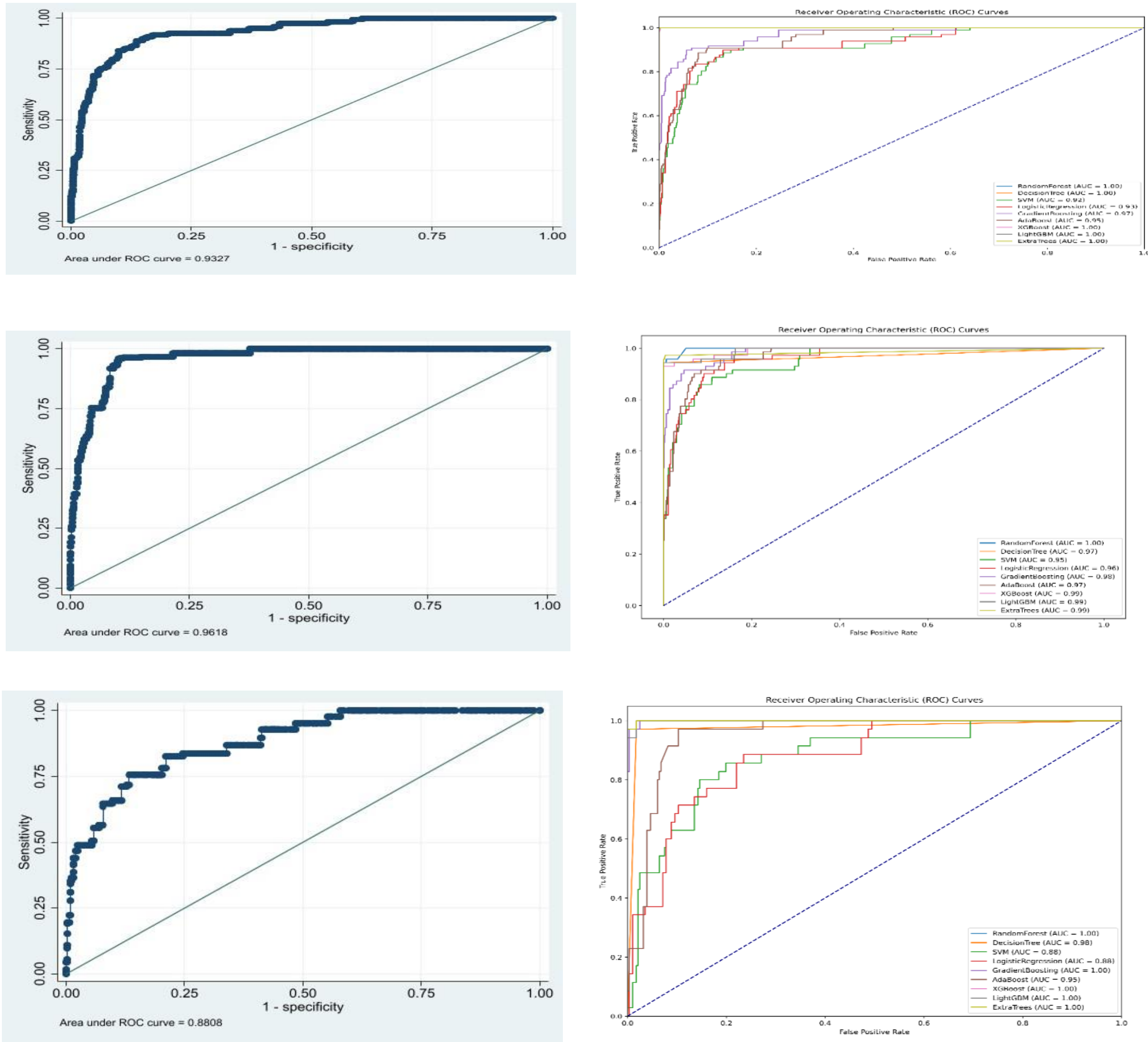
Appendix 4. Figure shows cross-validation accuracy and standard deviation for machine learning models applied to Roma settlements in Serbia, North Macedonia, and Montenegro. Random Forest, Extra Trees, and Light GBM are all machine learning ensemble models, and they demonstrate consistently high accuracy, and standard deviations that are almost zero, in all three countries, indicating a steady and reliable approach to poverty prediction. Some of their effectiveness in handling complex, multi-dimensional poverty data, is consistent with prior findings that such models feature intricate socio-economic relationships (Ferrer, & García, 2023). By contrast, SVC and Logistic Regression have lower accuracy and higher standard deviations, especially in Serbia and North Macedonia, and thus comparatively may be less capable of modeling the non-linear relations typically encountered in poverty-related data. The low variability and high accuracy of these models make them suitable for targeting resource-based poverty alleviation strategies in the WEFE Nexus context to address some key SDGs by identifying the resources along the WEFE Nexus (El-Meligi et al., 2024, Rogner, Bauer, 2022; UNECE 2022).



Appendix Table 2. Comparative Model Accuracy Analysis with Confidence Intervals across Roma Settlements in Serbia, North Macedonia, and Montenegro

Appendix Table 2. presents the test accuracy, mean accuracy, and the cross-validation (CV) standard deviation for different machine learning models applied to Roma settlements in Serbia, North Macedonia, and Montenegro. On the other hand, Ensemble models do well with Random Forest, XG Boost, and Extra Trees having high test and mean accuracies close to 0.99 and low CV standard deviations across all three countries. This stability is due to this strong predictive power as well as reliability, especially in their ability to pick up the esoteric, multi-dimensional facets of poverty within marginalized communities. Results from these examples demonstrate the utility of ensemble methods in the WEFE Nexus, which are useful for identifying critical resources, including water, energy, and sanitation as determinants of poverty. As these models increase the accuracy and consistency of predicting poverty, they are appropriate for SDGs, serving as robust tools to guide targeted poverty alleviation interventions (El-Meligi., et al., 2024, Rogner, & Bauer, 2022; UNECE, 2022).

Appendix Figure 3.



Appendix Table 2.

Model	Serbia (Roma Settlements)			N. Macedonia (Roma Settlements)			Montenegro (Roma Settlements)		
	Test Accuracy	Mean Accuracy	CV Std Dev	Test Accuracy	Mean Accuracy	CV Std Dev	Test Accuracy	Mean Accuracy	CV Std Dev
Random Forest	0.998547	0.996	0.002343	0.995079	0.997784	0.000922	0.996835	0.992879	0.002951
Decision Tree	0.998547	0.994	0.003796	0.996063	0.995568	0.002148	0.981013	0.987349	0.012346
SVM	0.941134	0.931636	0.007527	0.951772	0.947316	0.004012	0.889241	0.895483	0.009275
Logistic Regression	0.946221	0.933273	0.003618	0.951772	0.94707	0.005932	0.908228	0.902601	0.00969
Gradient Boosting	0.962209	0.963818	0.003118	0.969488	0.974399	0.006476	0.974684	0.97071	0.010474
Ada Boost	0.946948	0.938	0.003916	0.950787	0.950516	0.006307	0.889241	0.907353	0.009352
XG Boost	0.998547	0.995091	0.002545	0.995079	0.9968	0.001256	0.981013	0.991292	0.004613
Light GBM	0.998547	0.994182	0.002197	0.995079	0.996553	0.000922	0.981013	0.992882	0.005242
Extra Trees	0.998547	0.994545	0.003303	0.996063	0.997784	0.000922	0.996835	0.990501	0.003164

Appendix Figure 5