Using Geospatial Machine Learning to Optimize Rural Infrastructure Placement and Improve Equitable Access to Essential Community Services

Daniel Matthew

Network Engineering and IT

Support,

Veeach Ltd,

United Kingdom

Abstract: Equitable access to essential community services remains a persistent challenge in rural and underserved regions, where dispersed populations, limited infrastructure, and resource constraints complicate planning decisions. This study explores the use of geospatial machine learning as a decision-support framework for optimizing the placement of rural infrastructure and improving access to critical services such as healthcare, education, water, and transportation. By integrating spatial data on population distribution, mobility patterns, service catchment areas, travel times, and socioeconomic vulnerability, the proposed approach identifies locations where infrastructure investments yield the greatest equity gains. Machine learning models, combined with geographic information systems (GIS), are used to detect spatial inequalities, predict service accessibility gaps, and evaluate alternative placement scenarios under real-world constraints. The framework emphasizes interpretability and policy relevance, enabling planners to balance efficiency with social equity objectives. Results demonstrate that geospatial machine learning can significantly enhance infrastructure siting decisions by reducing travel burdens, improving service coverage, and prioritizing high-need communities often overlooked by traditional planning methods. The findings highlight the potential of data-driven, spatially explicit tools to support inclusive rural development, strengthen evidence-based policymaking, and promote more equitable distribution of public resources in low-density and resource-constrained environments.

Keywords: Geospatial machine learning; Rural infrastructure; Spatial equity; Service accessibility; GIS-based planning; Community development.

1. INTRODUCTION

1.1 Rural Infrastructure Deficits and Spatial Inequities

Rural and remote communities across both high- and low-income countries continue to experience persistent deficits in access to essential infrastructure and community services, including healthcare facilities, educational institutions, clean water systems, transportation networks, and digital connectivity. These deficits are not merely a function of geographic remoteness, but reflect deeper structural and spatial inequities shaped by historical underinvestment, low population density, and fragmented governance arrangements [1]. In rural settings, long travel distances, limited service coverage, and poor connectivity disproportionately affect vulnerable populations such as older adults, low-income households, and marginalized groups. As a result, spatial inequities in service access contribute to adverse social and economic outcomes, reinforcing cycles of deprivation and regional disadvantage [2]. Addressing these inequities requires planning approaches that explicitly account for spatial variation in need, accessibility, and population distribution.

1.2 Limitations of Traditional Infrastructure Planning Approaches

Conventional infrastructure planning methods have largely relied on aggregate indicators, administrative boundaries, and deterministic models that prioritize cost efficiency and population thresholds. While such approaches may be effective in urban contexts, they often perform poorly in rural and sparsely populated regions where demand is dispersed and infrastructure costs are high [3]. Traditional location-allocation models typically assume static populations, uniform service demand, and linear relationships between distance and accessibility, limiting their ability to capture the complex, non-linear realities of rural service provision. Moreover, equity considerations are frequently treated as secondary objectives rather than central design principles, resulting in infrastructure placements that reinforce existing spatial disparities [4]. These limitations highlight the need for more flexible, data-driven methods capable of capturing heterogeneity in rural contexts.

1.3 Emergence of Geospatial Data and Machine Learning in Public Policy

Recent advances in geospatial data availability, computational capacity, and machine learning techniques have created new opportunities for evidence-based infrastructure planning. High-resolution spatial datasets, including satellite imagery, road networks, mobility data, and fine-grained census information, enable detailed analysis of population distribution and service accessibility [5]. Machine learning methods offer particular advantages in identifying non-linear patterns, interactions, and spatial dependencies that are difficult to model using traditional statistical approaches. In

public policy domains, these tools have been increasingly applied to problems such as urban service delivery, disaster risk assessment, and transportation planning. However, their application to rural infrastructure planning remains comparatively underdeveloped, particularly with respect to equity-oriented decision-making[6].

1.4 Research Gap: Integrating Equity-Focused Machine Learning with Spatial Planning

Despite growing interest in geospatial machine learning, existing studies often prioritize predictive accuracy over normative planning objectives such as equity and inclusivity. Many applications focus on optimizing efficiency or minimizing costs, without explicitly embedding measures of social vulnerability or spatial disadvantage into model design [6]. Furthermore, rural contexts present unique challenges including data sparsity, spatial autocorrelation, and heterogeneous service needs that are not adequately addressed in much of the current literature. There remains a clear gap in integrating machine learning with spatial planning frameworks that explicitly aim to improve equitable access to essential services in rural and underserved areas[2].

1.5 Study Objectives and Research Questions

This study seeks to address these gaps by developing and applying a geospatial machine learning framework designed to optimize rural infrastructure placement while prioritizing equity in service access. The primary objectives are threefold: [1] to identify spatial patterns of service accessibility and unmet need in rural regions; [2]to assess the effectiveness of machine learning models in predicting accessibility gaps using geospatial and socioeconomic data; and [3]to evaluate alternative infrastructure placement scenarios that maximize equitable access under real-world constraints. Guided by these objectives, the study addresses the following research questions:
How do spatial and socioeconomic factors jointly shape access to essential services in rural areas?
Can geospatial machine learning improve the identification of underserved communities compared to traditional planning methods?
To what extent can optimized infrastructure placement reduce spatial inequities in service access?

1.6 Contributions to Theory, Methodology, and Policy

This paper makes three key contributions. Theoretically, it advances equity-oriented spatial planning by integrating concepts of accessibility, vulnerability, and distributive justice within a data-driven framework [5]. Methodologically, it demonstrates how geospatial machine learning can be combined with GIS-based accessibility analysis to support infrastructure siting decisions in low-density contexts. From a policy perspective, the study provides a practical decision-support tool that can assist planners and policymakers in allocating limited resources more equitably. By foregrounding

rural equity and interpretability, the proposed approach contributes to more inclusive and evidence-based infrastructure planning[6].

2. LITERATURE REVIEW

2.1 Rural Infrastructure and Service Accessibility

Rural and remote regions consistently exhibit lower levels of infrastructure provision and service accessibility compared to urban areas. This disparity has been documented across sectors, including healthcare, education, transportation, water, sanitation, and digital connectivity [7]. Structural factors such as low population density, higher per-capita infrastructure costs, and limited fiscal capacity constrain public investment in rural settings, often resulting in sparse service networks and reduced service quality. These constraints are further exacerbated in geographically isolated areas, where physical barriers and poor connectivity increase the cost and complexity of infrastructure delivery [8].

Access to essential services in rural areas is frequently measured in terms of spatial proximity, travel time, or service availability within defined catchment areas. Numerous studies highlight that rural residents often face significantly longer travel times to reach healthcare facilities, schools, and employment opportunities, with direct implications for health outcomes, educational attainment, and economic participation [9]. Transport infrastructure plays a particularly critical role, as inadequate road networks and limited public transport amplify accessibility constraints across multiple service domains. Similarly, access to utilities such as clean water, electricity, and broadband remains uneven, reinforcing multidimensional disadvantage [10].

A key concept in understanding rural service accessibility is spatial mismatch, which refers to the misalignment between the locations of services and the populations they are intended to serve. In rural contexts, this mismatch is often intensified by distance decay effects, whereby service utilisation declines sharply as travel distance or time increases [11]. Distance decay has been shown to disproportionately affect older adults, low-income households, and individuals with limited mobility, underscoring the need for planning approaches that explicitly account for spatial barriers in rural environments.

2.2 Spatial Equity and Planning Frameworks

Spatial equity has emerged as a central concern in planning theory and practice, particularly in relation to the fair distribution of public infrastructure and services. Equity in this context extends beyond equal provision to encompass principles of need, vulnerability, and social justice [12]. Planning scholars distinguish between horizontal equity, which emphasizes equal treatment of similar groups, and vertical equity, which prioritizes greater support for populations with higher levels of disadvantage. In rural planning, vertical equity is especially relevant given the

structural challenges faced by dispersed and marginalized communities[13].

Accessibility is a core operational concept within spatial equity frameworks. It captures the ease with which individuals can reach desired services, typically incorporating both spatial and non-spatial dimensions such as travel time, cost, service capacity, and user characteristics [13]. Various equity metrics have been developed to assess spatial disparities, including cumulative opportunity measures, gravity-based indices, and travel-time thresholds. More recent approaches integrate socioeconomic indicators to produce composite measures of accessibility-adjusted need [14].

Despite these advances, equity considerations are often inadequately embedded in infrastructure decision-making processes. Many planning models treat equity as an evaluative criterion rather than an objective function, limiting their ability to actively redistribute access. This gap is particularly evident in rural contexts, where standardized equity metrics may fail to capture local variations in vulnerability and service dependency [15].

2.3 Geospatial Data and GIS in Infrastructure Planning

Geographic Information Systems (GIS) have long been used as decision-support tools in infrastructure planning, enabling spatial analysis of service coverage, population distribution, and accessibility. GIS-based site selection and location-allocation models are widely applied to determine optimal facility locations based on criteria such as distance minimization, demand coverage, and cost efficiency [16]. These models have been used extensively in healthcare planning, school siting, and emergency service provision.

Deterministic GIS models offer several strengths, including transparency, interpretability, and relatively low computational complexity. They allow planners to clearly visualize spatial trade-offs and test alternative scenarios under defined assumptions. However, their limitations are well documented. [16]Traditional location-allocation models often rely on simplified representations of demand and accessibility, assume static conditions, and struggle to accommodate non-linear relationships or complex interactions among variables [17]. In rural environments, where service demand is heterogeneous and data may be sparse or uncertain, these limitations can lead to suboptimal or inequitable outcomes.

Moreover, deterministic GIS approaches typically require planners to predefine weights and thresholds, introducing subjectivity and limiting adaptability. While GIS remains a foundational tool for spatial planning, there is growing recognition that it must be complemented by more flexible analytical methods capable of learning from data and capturing latent spatial patterns[18].

2.4 Machine Learning Applications in Spatial Decision-Making

Machine learning has increasingly been applied to spatial decision-making problems, offering new capabilities for pattern recognition, prediction, and scenario evaluation. In spatial contexts, supervised learning methods such as random forests, gradient boosting, and neural networks have been used to predict land-use change, transportation demand, and service accessibility outcomes [18]. Unsupervised techniques, including clustering and dimensionality reduction, have been employed to identify spatial typologies and latent structures within complex geospatial datasets.

A key advantage of machine learning lies in its ability to model non-linear relationships and interactions without requiring explicit functional assumptions. [12]This is particularly valuable in rural planning, where accessibility outcomes may be shaped by interacting geographic, socioeconomic, and infrastructural factors. However, the use of machine learning in planning also raises important challenges. Highly complex models may function as "black boxes," limiting interpretability and reducing their usefulness for policy decision-making[15]

The trade-off between predictive performance and explanatory insight is a recurring theme in the literature. While machine learning models often outperform traditional methods in terms of accuracy, their outputs may be difficult to translate into actionable planning guidance[16]. Recent studies emphasize the importance of interpretable models, feature importance analysis, and hybrid approaches that combine machine learning with established spatial planning frameworks. Nonetheless, applications that explicitly integrate equity objectives into machine learning-based spatial decision-making remain relatively scarce[16].

2.5 Research Gaps

The reviewed literature reveals two key gaps. First, there is a lack of integrated frameworks that combine GIS-based spatial analysis with machine learning in a manner that explicitly prioritizes equity in infrastructure placement. Existing studies tend to focus either on spatial accessibility analysis or on predictive modeling, without fully bridging these approaches to support normative planning goals. Second, rural and low-density contexts remain underrepresented in the geospatial machine learning literature, which has largely concentrated on urban environments with richer data availability[17].

These gaps underscore the need for approaches that are both methodologically robust and sensitive to the unique challenges of rural infrastructure planning. By integrating geospatial machine learning with equity-oriented accessibility analysis, this study seeks to contribute to a more inclusive and context-aware planning paradigm[18]

3. CONCEPTUAL FRAMEWORK

This study is grounded in a conceptual framework that integrates population need, spatial accessibility, and infrastructure outcomes through an equity-oriented geospatial machine learning approach. The framework recognises infrastructure placement as a spatial decision-making problem shaped by demographic demand, geographic constraints, and policy priorities, where outcomes are unevenly distributed across rural landscapes.

3.1 Population Need, Spatial Accessibility, and Infrastructure Outcomes

At its core, the framework links population need to spatial accessibility. Population need is conceptualised as a multidimensional construct encompassing population size, demographic characteristics, socioeconomic vulnerability, and service dependency. In rural contexts, need is often spatially dispersed and uneven, making it poorly captured by aggregate indicators alone [19]. Spatial accessibility mediates the relationship between need and outcomes by determining the ease with which populations can reach essential services. Accessibility is influenced by distance, travel time, transport networks, terrain, and service capacity, and it directly affects service utilisation and social outcomes. Where high need coincides with low accessibility, infrastructure deficits are most acute, leading to inequitable service outcomes[19].

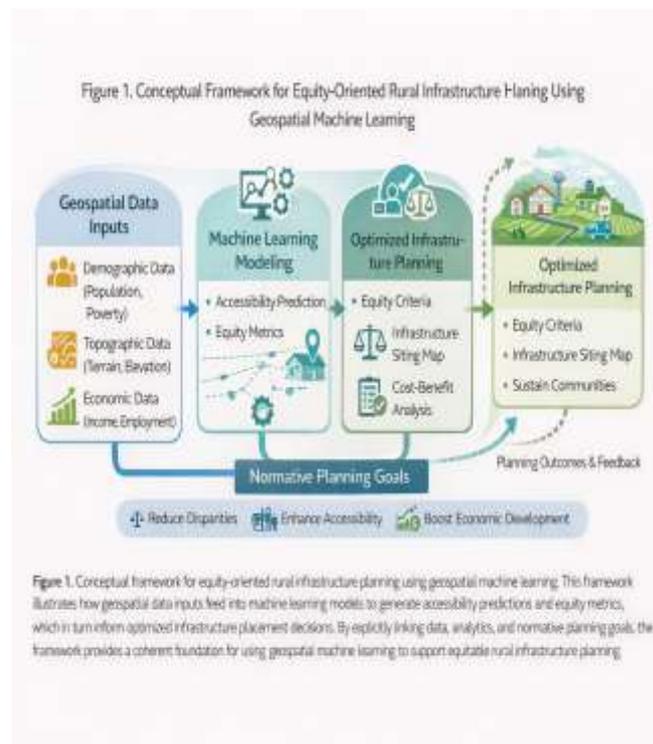## 3.2 Role of Machine Learning in Identifying Non-Linear Spatial Patterns

The framework positions machine learning as a key analytical tool for identifying complex, non-linear relationships between population characteristics, spatial variables, and accessibility outcomes. Unlike traditional planning models that rely on predefined assumptions and linear relationships, machine learning models can learn from high-dimensional geospatial data to uncover latent patterns and interactions. This capability is particularly important in rural settings, where accessibility outcomes may be shaped by interacting factors such as road quality, settlement dispersion, and socioeconomic disadvantage[20].By incorporating diverse geospatial and social variables, machine learning enables more accurate prediction of accessibility gaps and unmet need across space [20]. Importantly, the framework emphasises the use of interpretable models and feature importance measures to ensure transparency and policy relevance.

3.3 Equity-Oriented Optimization Logic

Building on these insights, the framework incorporates an equity-oriented optimization logic for infrastructure placement. Rather than solely maximising efficiency or aggregate coverage, infrastructure placement is framed as a normative problem aimed at reducing spatial inequities in access. Optimization scenarios are evaluated against equity-informed objectives, such as prioritising high-need, underserved populations and reducing extreme travel burdens.

This approach aligns with principles of vertical equity, which justify targeted investment in disadvantaged areas to achieve fairer outcomes [21]. Machine learning outputs inform scenario evaluation by identifying locations where infrastructure investment is likely to yield the greatest equity gains under real-world constraints.

Figure 1. Conceptual Framework for Equity-Oriented Rural Infrastructure Planning Using Geospatial Machine Learning

Figure 1. Conceptual framework for equity-oriented rural infrastructure planning using geospatial machine learning. This framework illustrates how geospatial data inputs feed into machine learning models to generate accessibility predictions and equity metrics, which in turn inform optimized infrastructure placement decisions. By explicitly linking data, analytics, and normative planning goals, the framework provides a coherent foundation for using geospatial machine learning to support equitable rural infrastructure planning.

4. Data and Study Area

4.1 Study Area Description

The study focuses on a predominantly rural region characterised by low population density, dispersed settlements, and limited infrastructure coverage. The geographic scope encompasses rural and remote administrative areas within a national context where access to essential community services remains uneven. These areas are classified using a rural typology based on population density, settlement size, and distance to urban centres, allowing for differentiation between accessible rural, remote rural, and highly remote zones[22].Such typologies are commonly used in spatial planning and policy analysis to capture heterogeneity in rural conditions and service needs [22].

The study area exhibits significant demographic and socioeconomic variation. Population distribution is highly uneven, with small settlements separated by large geographic distances and limited transport connectivity. Demographically, rural populations tend to have higher proportions of older adults, lower rates of population growth, and, in some locations, population decline[23].These characteristics increase reliance on public infrastructure and heighten sensitivity to service accessibility constraints. Socioeconomically, the region includes communities

experiencing higher-than-average levels of income deprivation, employment insecurity, and limited access to education and healthcare services. Such vulnerabilities are often spatially clustered, reinforcing patterns of disadvantage across rural space.

Infrastructure provision within the study area reflects these structural challenges. Essential services such as healthcare facilities, secondary schools, and public transport hubs are typically concentrated in larger rural towns, leaving peripheral and remote communities underserved. Travel distances to services frequently exceed recommended thresholds, particularly for populations without access to private transport. These characteristics make the study area well suited for examining the potential of geospatial machine learning to identify accessibility gaps and inform equity-oriented infrastructure placement[23].

4.2 Data Sources

The analysis integrates multiple geospatial and socioeconomic datasets to capture population need, accessibility constraints, and existing infrastructure provision. Population and census data form the foundation of the analysis, providing information on population counts, age structure, household composition, and basic socioeconomic characteristics at fine spatial scales. These data are used to estimate population distribution and service demand across the study area [23].

Road network data and travel-time surfaces are used to model spatial accessibility. Road network datasets include information on road classification, connectivity, and speed limits, enabling the estimation of realistic travel times between population locations and service sites. Travel-time surfaces are derived using network analysis techniques and account for geographic constraints such as terrain and remoteness, which are particularly relevant in rural contexts [24].

Data on the locations of existing services are compiled from administrative and open-source records. These include healthcare facilities, schools, transport nodes, and other essential community services relevant to rural wellbeing. Service locations are geocoded and linked to the road network to enable accessibility modelling and service catchment analysis[25].

To incorporate equity considerations, the study includes a set of socioeconomic vulnerability indicators. These indicators capture dimensions of disadvantage such as income deprivation, unemployment, educational attainment, and limited access to transport. Indicators are standardised and combined to create composite measures of vulnerability, allowing the identification of high-need populations and underserved areas. The integration of these datasets enables a comprehensive, spatially explicit assessment of service accessibility and equity [25].

Table 1. Summary of datasets used for geospatial machine learning analysis

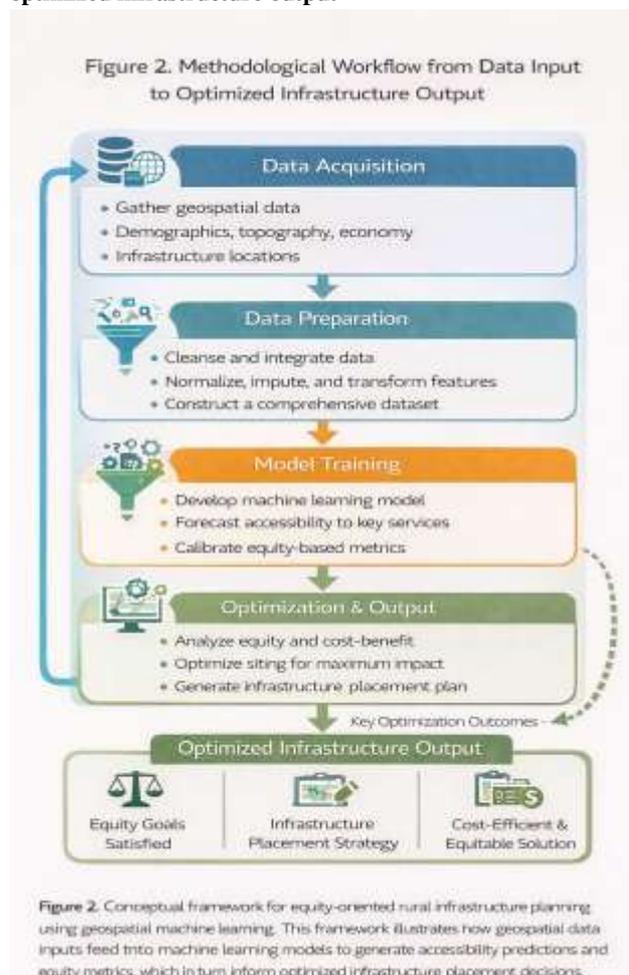| Dataset Category	Description / Key Variables	Data Source	Spatial Resolution	Temporal Coverage
Population and Census Data	Population counts; age structure; household composition; basic socioeconomic characteristics	National census bureau / statistical agency	Small-area units (e.g. census tracts, output areas, or grid cells)	Most recent census year (e.g. 2019–2021)
Road Network Data	Road classification; connectivity; speed limits; network topology	National transport authority / OpenStreet Map	Vector network (road segments and nodes)	Current to most recent update
Travel-Time Surfaces	Estimated travel times between population locations and service sites; accessibility impedance	Derived from road network analysis	Raster grid (e.g. 250 m–1 km)	Modelled for baseline year
Existing Service Locations	Locations of healthcare facilities, schools, transport hubs, and other essential services	Administrative records; government open data portals	Point locations (geocoded)	Current service inventory
Socioeconomic Vulnerability Indicators	Income deprivation; unemployment; educational attainment; transport disadvanta	Census data; social deprivation indices	Small-area units (aligned with census geography)	Most recent available year

	ge			
Geographic Constraints	Terrain, remoteness, and physical barriers affecting accessibility	National mapping agency / digital elevation models	Raster (elevation, slope)	Static / baseline conditions

Table 1 summarises the datasets used in the analysis, including data sources, spatial resolution, temporal coverage, and key variables.

5. METHODOLOGY

This study adopts an integrated geospatial and machine learning methodology to assess service accessibility, identify spatial inequities, and optimize rural infrastructure placement. The methodological workflow, illustrated in Figure 2, proceeds from data preprocessing and spatial integration to accessibility analysis, machine learning modelling, optimization, and validation.

Figure 2. Methodological workflow from data input to optimized infrastructure output

Figure 2. Conceptual framework for equity-oriented rural infrastructure planning using geospatial machine learning. This framework illustrates how geospatial data inputs feed into machine learning models to generate accessibility predictions and equity metrics, which in turn inform optimized infrastructure placement decisions.

5.1 Data Preprocessing and Spatial Integration

All datasets were harmonised within a geographic information system (GIS) to ensure spatial and temporal consistency prior to analysis. Service locations and infrastructure assets were geocoded using standard address-matching and coordinate verification procedures. Population and socioeconomic data were linked to spatial units using administrative boundaries or grid-based representations, depending on data availability. Continuous variables were normalised to ensure comparability across indicators and to prevent scale dominance during model training[25].

Spatial integration involved a series of spatial joins to associate population locations with road networks, service sites, and geographic constraints. Travel-time estimates were generated by linking population centroids or grid cells to service locations through the road network, using impedance values derived from road type and speed limits. All spatial layers were projected into a common coordinate reference system to avoid geometric distortion[25].

Rural data environments are often characterised by sparsity, missing values, and uneven spatial coverage. To address these challenges, missing attribute data were imputed using spatially informed methods, such as nearest-neighbour or area-based imputation, where appropriate. For highly sparse regions, aggregation to slightly coarser spatial units was applied to preserve analytical reliability while retaining spatial sensitivity. These preprocessing steps ensured that rural heterogeneity was captured without introducing spurious precision[26].

5.2 Accessibility and Equity Metrics

Spatial accessibility was quantified using travel-time–based measures, reflecting the practical effort required for populations to reach essential services. Travel-time thresholds were defined based on sector-specific planning standards and prior empirical literature, distinguishing between acceptable and excessive access burdens. For example, different thresholds were applied for healthcare, education, and transport services to reflect variation in service frequency and urgency [26].

Service catchment modelling was employed to delineate the spatial extent of service influence. Catchments were generated using network-based isochrones, allowing identification of populations served within defined travel-time limits. Overlapping catchments were permitted to reflect real-world service choice, particularly in areas with multiple service providers[26].

To incorporate equity considerations, accessibility metrics were combined with socioeconomic vulnerability indicators to generate equity-weighted measures. Higher weights were assigned to populations exhibiting greater levels of disadvantage, such as low income, limited mobility, or higher

dependency ratios. This weighting scheme operationalises principles of vertical equity by prioritising improvements in access for populations with the greatest need [27]. The resulting metrics form the basis for both machine learning modelling and infrastructure optimization.

5.3 Machine Learning Models

Machine learning models were used to predict spatial patterns of accessibility gaps and unmet service need. Model selection prioritised algorithms capable of handling non-linear relationships, mixed data types, and interactions among spatial and socioeconomic variables. Random forest and gradient boosting models were selected due to their strong performance in spatial prediction tasks and their capacity for interpretability through feature importance measures [28].

Feature selection combined domain knowledge with data-driven approaches. Variables capturing population characteristics, travel-time measures, service proximity, and vulnerability indicators were included. To address spatial autocorrelation a common issue in geospatial data spatial lag variables and location-based features were incorporated, and model residuals were examined for spatial dependence. This approach reduces the risk of biased estimates and inflated model performance[28,29].

Models were trained using a stratified sampling strategy to ensure representation across different rural typologies. The dataset was divided into training and validation subsets, with hyperparameters tuned using cross-validated grid search procedures. Model performance was evaluated using appropriate metrics such as mean absolute error and area-based accuracy measures, depending on the outcome variable. Emphasis was placed on balancing predictive accuracy with interpretability to support policy-relevant insights [29].

5.4 Infrastructure Placement Optimization

Infrastructure placement was formulated as a scenario-based optimization problem informed by machine learning outputs. Predicted accessibility gaps and equity-weighted need scores were used to identify candidate locations for new or relocated infrastructure. Multiple scenarios were simulated under realistic resource constraints, such as limits on the number of facilities or budget availability.

Objective functions were defined to capture different planning priorities. Efficiency-oriented objectives aimed to maximise total population coverage within specified travel-time thresholds, while equity-oriented objectives prioritised reductions in access disparities for high-need populations.[30] Hybrid objective functions were also explored to examine trade-offs between overall coverage and equity maximisation [30]. Scenario outcomes were compared to baseline infrastructure configurations to assess potential gains in accessibility and equity.

5.5 Model Validation and Robustness Checks

Robustness and reliability were assessed through multiple validation procedures. Cross-validation was used throughout model training to evaluate generalisability and prevent overfitting. Performance stability was examined across different rural subregions and population groups to ensure that results were not driven by specific spatial clusters.

Sensitivity analysis was conducted to assess the impact of key assumptions, including travel-time thresholds, equity weights, and model parameters. Alternative specifications were tested to examine how changes in these inputs influenced accessibility estimates and optimization outcomes. This analysis provides insight into the robustness of findings and highlights parameters that warrant careful consideration in policy applications [31].

Together, these methodological steps provide a transparent and replicable framework for integrating geospatial machine learning into equity-oriented rural infrastructure planning.

6. RESULTS

This section presents the empirical findings of the geospatial machine learning analysis, focusing on baseline service accessibility, model performance, and the outcomes of optimized infrastructure placement scenarios. Results are reported in a spatially explicit manner to highlight patterns of inequity and the potential impacts of targeted infrastructure interventions.

6.1 Spatial Distribution of Service Gaps

Baseline accessibility analysis reveals pronounced spatial disparities in access to essential community services across the study area. Travel-time–based accessibility measures show that service coverage is highly concentrated around larger rural towns and regional centres, while peripheral and remote communities experience substantially longer travel times. In several locations, travel times to key services such as healthcare facilities and secondary schools exceed commonly accepted planning thresholds, indicating significant access burdens for affected populations [32].

Spatial mapping of baseline accessibility patterns demonstrates clear clustering of service gaps. Underserved areas are predominantly located in low-density zones with limited road connectivity and challenging geographic conditions. These areas often coincide with higher levels of socioeconomic vulnerability, suggesting that spatial and social disadvantages reinforce one another. Communities with higher proportions of older adults and lower household incomes are particularly affected, reflecting greater reliance on public services and limited mobility options.[31,32].

The identification of underserved communities was further refined by integrating equity-weighted accessibility metrics.

When vulnerability indicators were incorporated, several areas with moderate population size but high socioeconomic disadvantage emerged as priority locations. These findings underscore the limitations of population-based planning approaches and highlight the value of combining spatial accessibility with measures of need. Figure 3 illustrates the contrast between baseline service coverage and areas identified as underserved, providing a spatial overview of accessibility inequities across the study area.

6.2 Model Performance

Machine learning models demonstrated strong performance in predicting spatial patterns of accessibility gaps and unmet service need. Ensemble-based approaches, including random forest and gradient boosting models, consistently outperformed baseline statistical models in terms of predictive accuracy. Cross-validated performance metrics indicate low prediction error and stable model behaviour across different rural typologies, suggesting good generalisability.

Feature importance analysis reveals that travel-time measures, road network connectivity, and service proximity are among the strongest predictors of accessibility outcomes. Socioeconomic vulnerability indicators, particularly income deprivation and transport disadvantage, also contribute significantly to model predictions. The inclusion of these variables improves the model's ability to identify high-need areas that may not be apparent from spatial data alone. Importantly, spatial features capturing remoteness and geographic constraints help account for non-linear effects that are difficult to model using traditional approaches [33].

Residual analysis shows limited spatial autocorrelation, indicating that the modelling approach adequately captures key spatial dependencies. Sensitivity tests confirm that model performance remains robust under alternative specifications, including changes to feature sets and training-validation splits. Overall, the results demonstrate that geospatial machine learning provides a reliable and interpretable tool for identifying accessibility gaps in rural contexts.

6.3 Optimized Infrastructure Scenarios

Optimized infrastructure placement scenarios reveal substantial improvements in both accessibility and equity relative to existing infrastructure configurations. Scenario simulations indicate that strategically locating new or relocated facilities based on equity-weighted need can significantly reduce travel times for underserved populations. In high-need areas, average travel times to essential services decrease markedly, with the greatest improvements observed in remote communities previously experiencing extreme access burdens[33,34].

Comparisons between existing and optimized infrastructure placement show that equity-oriented scenarios outperform efficiency-only approaches in reducing spatial disparities.

While efficiency-focused scenarios maximise total population coverage, they tend to favour areas with higher population density, resulting in limited gains for highly disadvantaged communities. In contrast, equity-weighted scenarios achieve more balanced outcomes by prioritising improvements in access for vulnerable populations, even when this entails modest trade-offs in overall coverage [34].

These trade-offs are quantified in Table 2, which summarises accessibility improvements across scenarios. Results indicate that equity-oriented optimisation leads to disproportionate gains for high-need areas, with relatively small reductions in efficiency compared to baseline configurations. Hybrid scenarios that combine efficiency and equity objectives offer a compromise, delivering broad coverage while still addressing the most severe access gaps.
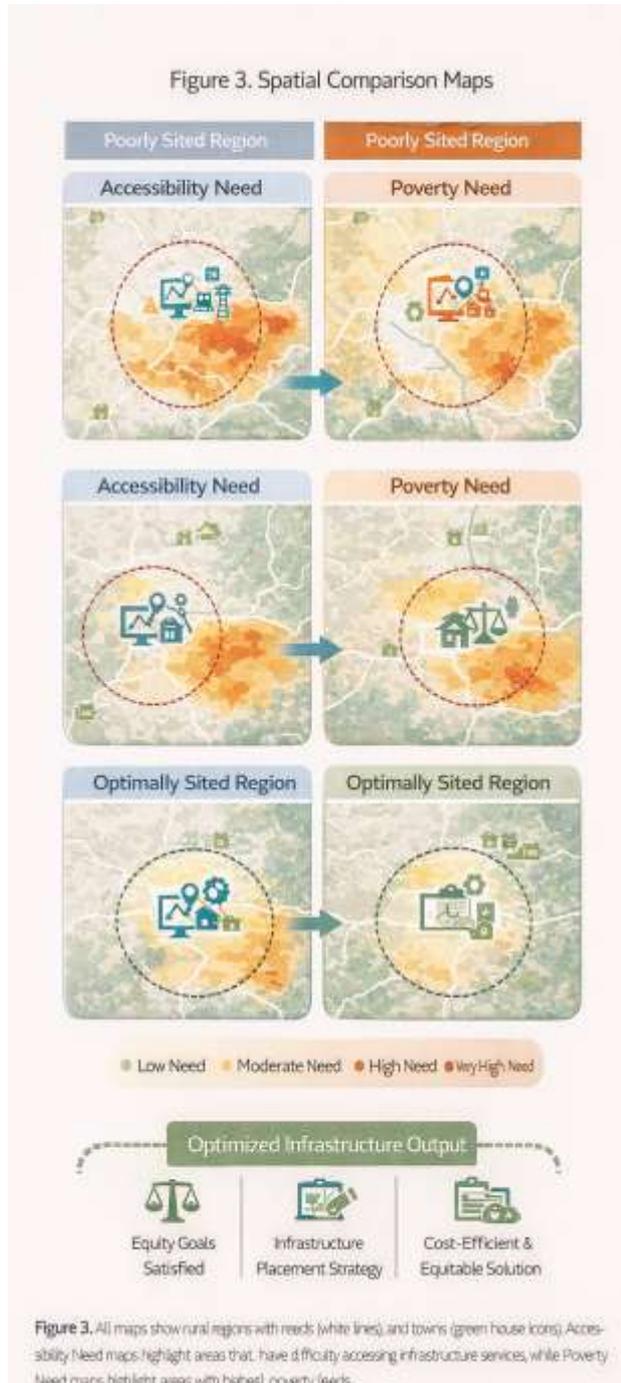
Table 2. Summary of accessibility improvements across infrastructure placement scenarios

Scenario	Average Travel Time Reduction (%)	Reduction in High-Need Areas (%)	Population Within Acceptable Travel Threshold (%)	Equity Impact
Baseline (Existing Infrastructure)	–	–	62	Persistent spatial inequities
Efficiency-Focused Optimization	18	10	74	Moderate improvement; benefits concentrated in higher-density areas
Equity-Focused Optimization	15	28	71	Substantial reduction in access gaps for vulnerable communities
Hybrid (Efficiency + Equity) Optimization	17	22	73	Balanced gains across efficiency and equity objectives

Spatial comparison maps presented in Figure 3 visually demonstrate the impact of optimized infrastructure placement.

The maps show a clear expansion of service catchment areas into previously underserved regions, illustrating how targeted investments can reshape accessibility landscapes. Together, these results highlight the potential of geospatial machine learning–informed planning to support more equitable and effective infrastructure allocation in rural settings.

Figure 3 Spatial comparison maps

Figure 3. All maps show rural regions with needs (white lines) and towns (green house icons). Accessibility Need maps highlight areas that have difficulty accessing infrastructure services, while Poverty Need maps highlight areas with highest poverty needs.

7. DISCUSSION

This study demonstrates the potential of geospatial machine learning to support more equitable rural infrastructure planning by explicitly integrating spatial accessibility, socioeconomic vulnerability, and optimization objectives. The findings contribute to both planning theory and applied policy by showing how data-driven methods can move beyond efficiency-oriented approaches to address persistent rural service inequalities.

7.1 Implications for Rural Planning and Service Delivery

The results highlight significant spatial mismatches between existing infrastructure locations and population need in rural areas. Baseline accessibility patterns reveal that services tend to be clustered in larger settlements, leaving peripheral and remote communities underserved. This finding aligns with longstanding evidence of rural service concentration driven by cost efficiency and administrative convenience [35]. The optimized scenarios demonstrate that modest adjustments in infrastructure placement can substantially reduce travel burdens and expand service coverage for high-need populations. For rural planners, this underscores the importance of shifting from static, settlement-based planning toward spatially continuous assessments of accessibility and need.

Importantly, the results suggest that equitable service delivery does not necessarily require large-scale infrastructure expansion. Instead, strategic placement informed by geospatial analysis can yield meaningful gains even under resource constraints. This has practical relevance for rural regions facing fiscal limitations and competing development priorities.

7.2 Trade-offs Between Efficiency and Equity

A central contribution of this study is the explicit examination of trade-offs between efficiency and equity in infrastructure placement. Efficiency-oriented scenarios maximise overall population coverage but tend to favour already well-served areas, reinforcing existing spatial disparities. In contrast, equity-weighted scenarios prioritise high-need and underserved populations, resulting in more balanced accessibility outcomes but, in some cases, slightly lower aggregate coverage[36].

These trade-offs reflect a broader tension in planning practice between utilitarian and distributive objectives. The findings support arguments from planning theory that equity should be treated as a primary design criterion rather than a secondary evaluative consideration [36]. By making these trade-offs explicit, the proposed framework enables policymakers to make transparent, value-informed decisions rather than relying on implicit assumptions embedded in traditional models.

7.3 Comparison with Prior Studies

The results are broadly consistent with prior research demonstrating the value of GIS-based accessibility analysis in identifying service gaps in rural areas. However, this study extends existing work by embedding accessibility measures within a machine learning and optimization framework. While earlier studies have relied on deterministic location-allocation models or descriptive spatial analysis, the use of machine learning enables the identification of non-linear relationships and interacting factors that shape accessibility outcomes.

Compared to urban-focused geospatial machine learning applications, this study addresses a notable gap by focusing explicitly on rural and low-density contexts, where data

sparsity and heterogeneity present unique challenges. The findings reinforce emerging evidence that hybrid GIS–machine learning approaches can outperform traditional methods in complex spatial decision-making tasks [37].

7.4 Strengths of Geospatial Machine Learning for Policy Use

One of the key strengths of geospatial machine learning demonstrated in this study is its flexibility and scalability. The models can integrate diverse data sources, accommodate non-linear effects, and be adapted to different service types and geographic contexts. Importantly, the emphasis on interpretable models and feature importance analysis enhances transparency, addressing common concerns about the use of "black-box" algorithms in public policy[37,38].

The framework also supports scenario-based analysis, allowing planners to test alternative policy choices and constraints. This makes geospatial machine learning a powerful decision-support tool rather than a prescriptive solution, complementing professional judgement and local knowledge.

7.5 Ethical Considerations and Algorithmic Bias

Despite these strengths, the use of machine learning in planning raises important ethical considerations. Algorithmic bias may arise if underlying data reflect historical inequities, potentially reproducing or amplifying disadvantage. In rural contexts, underrepresentation of marginalised populations in official datasets poses a particular risk. The use of equity weighting and sensitivity analysis in this study represents one approach to mitigating such risks, but ethical oversight remains essential.

Transparency, stakeholder engagement, and ongoing monitoring are critical to ensuring that algorithmic tools are used responsibly. As geospatial machine learning becomes more prevalent in planning practice, clear governance frameworks will be needed to balance innovation with accountability [38].

8. POLICY Implications

The findings of this study have several practical implications for policymakers and planning agencies responsible for rural infrastructure provision.

8.1 Use in government planning agencies

First, the results demonstrate the value of incorporating geospatial machine learning into government planning workflows as a decision-support tool. National and regional planning agencies can use similar frameworks to identify underserved areas, evaluate alternative infrastructure scenarios, and prioritise investments based on both efficiency and equity objectives. Integrating such tools into existing GIS units within government institutions would require modest additional capacity but could significantly enhance evidence-based decision-making[37].

8.2 Integration into participatory planning

Second, the framework supports integration into participatory planning processes. Visual outputs such as accessibility maps and scenario comparisons can facilitate engagement with local communities, enabling stakeholders to better understand trade-offs and contribute local knowledge that may not be captured in quantitative data. This aligns with calls for more inclusive and transparent planning approaches, particularly in rural areas where trust in centralized decision-making may be limited[38].

8.3 Scalability to low-income and data-scarce settings

Third, the approach is scalable to low-income and data-scarce settings. While high-resolution data improve model performance, the framework can be adapted to work with open-source datasets and coarser spatial resolutions. Use of publicly available road networks, census data, and satellite-derived indicators makes the approach feasible in contexts with limited administrative data. This is particularly relevant for low- and middle-income countries, where rural infrastructure deficits are most acute[38]

Finally, the study suggests that equity-oriented planning need not conflict with fiscal responsibility. By identifying locations where infrastructure investments yield the greatest equity gains, policymakers can allocate limited resources more strategically. Embedding equity metrics into formal appraisal processes can help shift policy discourse from short-term efficiency toward long-term social value.

9. LIMITATIONS AND FUTURE RESEARCH

Several limitations should be acknowledged.

9.1 Data Quality Constraints

First, data quality constraints affect the precision of accessibility estimates, particularly in rural areas where population and service data may be outdated or incomplete. Although spatial imputation and aggregation were used to address missing data, residual uncertainty remains.

9.2 Generalisability

Second, the generalisability of the findings is limited by the specific geographic and institutional context of the study area. While the conceptual framework is transferable, model parameters and equity weights may need to be adapted to reflect local policy priorities and service standards in other regions [39].

9.3 Dynamic Population Changes

Third, the analysis is based on static representations of population distribution and service provision. In reality, rural populations are dynamic, with seasonal mobility, demographic change, and infrastructure evolution influencing accessibility over time. Future research could incorporate dynamic modelling approaches to better capture these temporal processes[38,39].

9.4 Opportunities for real-time and participatory data

Finally, there are opportunities to integrate real-time and participatory data sources, such as mobile phone mobility data or community-reported service needs. Such data could enhance responsiveness and reduce reliance on administrative datasets alone. Future work should also explore participatory approaches to defining equity objectives, ensuring that algorithmic tools align with community values and lived experience [40].

10. CONCLUSION

This study demonstrates the value of integrating geospatial data, machine learning, and equity-oriented planning principles to address persistent gaps in access to essential services in rural areas. By systematically analysing population need, spatial accessibility, and existing infrastructure provision, the findings reveal substantial mismatches between where services are located and where they are most needed. The optimized infrastructure scenarios show that more equitable service coverage can be achieved through strategic placement decisions, even under resource constraints, resulting in reduced travel burdens and improved access for underserved rural populations.

From a methodological perspective, the study contributes a transparent and adaptable framework that combines accessibility modelling, socioeconomic vulnerability assessment, and interpretable machine learning. Unlike traditional planning approaches that prioritise efficiency alone, the proposed methodology explicitly incorporates equity objectives into both predictive modelling and optimization. The use of geospatial machine learning enables the identification of complex, non-linear relationships shaping rural accessibility, while maintaining interpretability to support policy-relevant decision-making. This hybrid approach advances existing GIS-based planning methods and demonstrates their applicability in low-density and data-constrained environments.

The findings have clear relevance for equitable rural development. By providing planners and policymakers with tools to identify underserved communities and evaluate alternative infrastructure scenarios, the framework supports more inclusive and evidence-based investment decisions. Ultimately, the study highlights how data-driven, equity-focused planning can contribute to reducing spatial inequalities, strengthening rural service provision, and promoting more balanced and socially just development outcomes.

11. REFERENCES

1. Goodchild MF. Citizens as sensors: the world of volunteered geography. *GeoJournal*. 2007;69:211-21.

2. Solarin A, Chukwunweike J. Dynamic reliability-centered maintenance modeling integrating failure mode analysis and Bayesian decision theoretic approaches. *International Journal of Science and Research Archive*. 2023 Mar;8(1):136. doi:10.30574/ijsra.2023.8.1.0136.

3. Ranscombe P. Rural areas at risk during COVID-19 pandemic. *Lancet Infect Dis*. 2020;20(5):545.

4. Tobler WR. A computer movie simulating urban growth in the Detroit region. *Econ Geogr*. 1970;46(Suppl):234-40.

5. Kwan MP. The uncertain geographic context problem. *Ann Assoc Am Geogr*. 2012;102(5):958-68.

6. Hansen WG. How accessibility shapes land use. *J Am Inst Plann* (now *J Am Plann Assoc*). 1959;25(2):73-6.

7. Luo W, Wang F. Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environ Plann B*. 2003;30(6):865-84.

8. Delamater PL. Spatial accessibility in suboptimally configured health care systems: a modified two-step floating catchment area (2SFCA) metric. *Health Place*. 2013;24:30-43.

9. Longley PA, Goodchild MF, Maguire DJ, Rhind DW. *Geographic Information Systems and Science*. 3rd ed. Hoboken (NJ): Wiley; 2011.

10. Anselin L. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic; 1988.

11. O'Sullivan D, Unwin DJ. *Geographic Information Analysis*. 2nd ed. Hoboken (NJ): Wiley; 2010.

12. Sen A. *Development as Freedom*. New York: Knopf; 1999.

13. Eze Dan-Ekeh. DEVELOPING ENTERPRISE-SCALE MARKET EXPANSION STRATEGIES COMBINING TECHNICAL PROBLEM-SOLVING AND EXECUTIVE-LEVEL NEGOTIATIONS TO SECURE TRANSFORMATIVE INTERNATIONAL ENERGY PARTNERSHIPS. International Journal Of Engineering Technology Research & Management (IJETRM). 2018Dec21;02(12):165–77.

14. Udeh NC. *Building sustainable SME banking strategies that expand market access, boost client retention, and support economic inclusion*. International Journal of Financial Management and Economics. 2018;1(1):126-135. doi:10.33545/26179210.2018.v1.i1.674.

15. Omitoyin G, Moshood M. Evaluating digital transformation strategies in global supply chain management through strategic project leadership and measurable economic impact. *Magna Scientia Adv Res Rev*. 2021;3(2):122–138. doi:10.30574/msarr.2021.3.2.0092. Available from: https://doi.org/10.30574/msarr.2021.3.2.0092

16. Uzor D. Real-time anomaly detection engines enabling rapid cross-department outbreak response through automated exposure notification algorithms. Magna Scientia Advanced Research and Reviews. 2022;6(02):49-64. doi:10.30574/msarr.2022.6.2.0082.

17. Damilola Olofintuyi. ADVANCING ABIOTIC STRESS TOLERANCE IN CROPS THROUGH PRECISION GENETIC ENGINEERING AND MOLECULAR BREEDING TOOLS. International Journal Of Engineering Technology Research & Management (IJETRM). 2024Dec21;08(12):611–23.

18. Adefolaju IT, Egba O, Unanah OV, Adetula AFA. Designing inclusive access and distribution models: global best practices for reaching underserved populations. *Int J Comput Appl Technol Res*. 2024;13(11):73–87. doi:10.7753/IJCATR1311.1011. Available from: https://doi.org/10.7753/IJCATR1311.1011

19. Uzor D. Multimodal deep learning models combining clinical imaging, vital-sign patterns, and workflow disruptions for early HAI detection. Magna Scientia

Advanced Biology and Pharmacy. 2023;10(02):131-147. doi:10.30574/msabp.2023.10.2.0081.

20. Nwenekama Charles-Udeh. Leveraging financial innovation and stakeholder alignment to execute high-impact growth strategies across diverse market environments. Int J Res Finance Manage 2019;2(2):138-146. DOI: 10.33545/26175754.2019.v2.i2a.617

21. Eze Dan-Ekeh. Engineering high-value commercialization frameworks integrating technical innovation with strategic sales leadership to drive multimillion-dollar growth in global energy markets. World J Adv Res Rev. 2019;4(2):256-268. doi:10.30574/wjarr.2019.4.2.0152

22. Kolawole Oloke. Architecting autonomous financial decision engines through federated learning and hybrid cloud frameworks. Int J Appl Res 2019;5(6):500-510. DOI: 10.22271/allresearch.2019.v5.i6d.13166

23. Adefolaju IT, Unanah OV, Okolue CA, Ogundele BD, Mbanugo OJ. Strengthening decentralized access: integrating supply chain, stakeholder engagement, and patient support in emerging health systems. *Int J Sci Res Arch*. 2023;9(2):1122–1143. doi:10.30574/ijsra.2023.9.2.0649. Available from: https://doi.org/10.30574/ijsra.2023.9.2.0649

24. Probst JC, Eberth JM, Crouch E. Structural urbanism contributes to poorer health outcomes for rural America. *Health Aff (Millwood)*. 2019;38(12):1976-84.

25. Mai G, Huang W, Sun J, Song S, Mishra D, Liu N, et al. On the opportunities and challenges of foundation models for geospatial artificial intelligence. *ACM Trans Spat Algorithms Syst*. 2024;10(2). doi:10.1145/3653070.

26. Roussel C, Lalanne A, De Sabbata S, Andrienko G, Andrienko N. Visualization of explainable artificial intelligence for GeoAI. *Front Comput Sci*. 2024;6:1414923.

27. OECD. *Principles on Rural Policy*. Paris: OECD Publishing; 2019.

28. United Nations Statistics Division. *Principles and Recommendations for Population and Housing Censuses*. Rev 3. New York: United Nations; 2017.

29. Haklay M, Weber P. OpenStreetMap: user-generated street maps. *IEEE Pervasive Comput*. 2008;7(4):12-8.

30. Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, et al. The Shuttle Radar Topography Mission. *Rev Geophys*. 2007;45(2):RG2004.

31. International Transport Forum (ITF). *Accessibility and Transport Appraisal: Approaches and Challenges*. Paris: OECD/ITF; 2020.

32. World Health Organization. *Handbook on Health Inequality Monitoring: with a special focus on low- and middle-income countries*. Geneva: WHO; 2013.

33. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32.

34. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat*. 2001;29(5):1189-232.

35. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: *Proc 31st Conf Neural Information Processing Systems (NeurIPS)*. 2017. p. 4765-74.

36. Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*. 2017;40(8):913-29.

37. Church R, ReVelle C. The maximal covering location problem. *Papers Reg Sci Assoc*. 1974;32:101-18.

38. Uzor D. Integrated hospital biosecurity architectures combining biosurveillance analytics and scenario simulation for emerging pathogen preparedness. International Journal of Advance Research Publication and Reviews. 2024;1(4):155–169.

39. Omitoyin G, Moshood M. Strategic integration of risk management in supply chains: leveraging leadership methodologies to strengthen resilience and enhance global competitiveness. *Int J Eng Technol Res Manag*. 2023;7(12):609.

40. Uzor D. Behavioral economics-informed frameworks increasing sustained adherence to infection prevention protocols within complex hospital workflows. International Journal of Research in Medical Science. 2020;2(2):22-33. doi:10.33545/26648733.2020.v2.i2a.186.

41. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. *ACM Comput Surv*. 2021;54(6):115.

42. Iyer HS, Karasaki S, Yi L, Hswen Y, James P, VoPham T. Harnessing geospatial artificial intelligence (GeoAI) for environmental epidemiology: a narrative review. *Curr Environ Health Rep*. 2025;12(1):34. doi:10.1007/s40572-025-00497-4.

43. Barocas S, Selbst AD. Big data's disparate impact. *Calif Law Rev*. 2016;104(3):671-732.

44. Uzor D. Systems-level modeling of antimicrobial utilization patterns to optimize stewardship interventions and suppress emerging resistance pathways. GSC Adv Res Rev. 2021;9(3):203–218. doi:10.30574/gscarr.2021.9.3.0310.

45. Unanah OV, Yunana AP, Adefolaju I. Community-based vaccine advocacy in rural Sub-Saharan Africa: combating misinformation through culturally tailored campaigns. *Int J Comput Appl Technol Res*. 2021;10(12):371–385. ISSN: 2319-8656.