MIJRD Multidisciplinary International Journal of Research and Development

Multidisciplinary International Journal of Research and Development

Volume: 04 / Issue: 06 / 2025 - Open Access - Website: www.mijrd.com - ISSN: 2583-0406

Explainable Multimodal Approach for Infrastructure-Driven Socioeconomic Assessment Using Satellite Imagery and Open-Source LLMs

Dibyendu Banerjee[1], Sourav Kairi[2], and Srijani Das[3]

[1]Sr. Manager, Cognizant Technology Solutions, Kolkata, West Bengal

[2,3]Associate, Cognizant Technology Solutions, Kolkata, West Bengal

Abstract— Accurate and interpretable assessment of socioeconomic conditions is vital for equitable policy-making, yet traditional methods face limitations in scalability and resolution. This paper presents a novel explainable multimodal framework that synergizes high-resolution satellite imagery with large language models (LLMs) to infer infrastructure- driven development levels. Unlike opaque deep learning models, our approach extracts visual features (e.g., roof materials, road density) and leverages LLMs to generate human-understandable insights, bridging the gap between geospatial data and actionable policy recommendations. We propose a hybrid architecture that aligns visual and textual modalities, enabling transparent analysis of underdeveloped regions. Experiments on diverse urban datasets demonstrate superior performance in identifying informal settlements and infrastructure gaps. By integrating explainability with remote sensing, this work advances responsible AI for social impact, offering governments and NGOs a scalable tool for targeted interventions.

Keywords— Explainable AI, Socioeconomic Assessment, Satellite Imagery, Large Language Models, Multimodal Learning, Urban Development, Remote Sensing, Infrastructure Mapping, Interpretable Machine Learning, Sustainable Development.

I. INTRODUCTION

Socioeconomic inequality and infrastructure disparity are persistent challenges in many regions, particularly within rapidly urbanizing areas of the Global South. Accurately identifying underdeveloped zones is crucial for effective urban planning, equitable policy-making, and targeted resource allocation. However, conventional approaches to socioeconomic assessment—such as census surveys, household interviews, and administrative reporting—are often limited by high operational costs, coarse spatial resolution, infrequent updates, and data accessibility constraints. These limitations hinder timely and localized decision-making, which is vital for dynamic and growing urban environments.[1]

Recent advances in remote sensing (RS) and artificial intelligence (AI) have created new opportunities to perform large-scale assessments of infrastructure and development patterns using satellite imagery. Satellite-derived features such as roof typologies, building densities, road networks, and land- use distributions offer valuable indirect indicators of socioeconomic status.[1] Yet, many existing machine learning applications in RS are black-box models, lacking transparency and interpretability—two essential qualities for socially impactful and trustworthy AI systems.[2] To address these challenges, this paper proposes an Explainable Multimodal Approach for Infrastructure-Driven Socioeconomic Assessment, which leverages high-resolution satellite imagery in

Volume: 04 / Issue: 06 / 2025 - Open Access - Website: www.mijrd.com - ISSN: 2583-0406

combination with large language models (LLMs). This approach aims to provide interpretable, data-driven insights into infrastructure quality and spatial development conditions that can be used as proxies for socioeconomic status.[5][6] By integrating visual feature extraction with natural language reasoning, the model produces human-understandable explanations, facilitating trust, accountability, and informed action by government agencies, NGOs, and urban planners.

Unlike traditional classification or segmentation models that focus purely on prediction, our framework prioritizes explainability and human-aligned interpretation. The method enables analysts to understand why certain urban regions are identified as underdeveloped by linking infrastructure patterns— such as irregular roof types or sparse construction grids—to likely socioeconomic indicators via LLM-generated narratives. This multimodal fusion architecture unites spatial visual features with semantic reasoning, offering a novel perspective on urban analysis

The primary contributions of this work are:

Proposing a multimodal hybrid architecture that integrates spatial features from high-resolution satellite imagery with the reasoning and interpretability capabilities of using LLMs effectively;

- Enabling human-readable explanations of infrastructure conditions to enhance the transparency of development assessments;
- Demonstrating the effectiveness of the approach on real-world urban datasets for identifying areas with insufficient or informal infrastructure;
- Bridging explainable AI and remote sensing, offering a scalable and socially responsible framework for infrastructure-aware socioeconomic evaluation.

This research advances the application of explainable multimodal AI in the field of urban analytics and remote sensing. By shifting focus from purely predictive outputs to interpretable insights, it facilitates data-informed governance and inclusive development planning.

II. ALGORITHM USED

This study introduces a modular, end-to-end framework that integrates geospatial image, prompt engineering, large language models (LLMs), and structured semantic parsing to produce explainable, multimodal socioeconomic assessments derived from satellite-inferred urban morphology. The proposed pipeline comprises seven interconnected components: [7] (a) geospatial prompt engineering,

(b) Google Gemini 2.5 Pro LLM, (c) semantic parsing and data normalization, (d) socioeconomic classification and typology extraction, (e) multivariate visualization and spatial diagnostics, (f) equity quantification using entropy and inequality metrics, and (g) explainability and observability mechanisms.

a. Geospatial Prompt Engineering

To ensure accurate, context-sensitive, and policy- relevant urban analysis, the framework integrates a carefully structured prompt engineering strategy that encodes role orientation, task-specific instructions, and output formatting guidance. This enhances the alignment between the LLM's generative capabilities and the analytical demands of geospatial classification.

- ***Role Specification:*** Prompts are designed to explicitly assign the LLM a domain-relevant role (here advanced geospatial analysis AI trained on extensive datasets), thereby calibrating the model's response to adopt appropriate tone, terminology, and analytical depth consistent with expert-level discourse. This sets the perspective or role from which the model should generate the content, ensuring the tone and depth of information are appropriate for an expert in the field.[13]
- ***Instruction:*** Each prompt embeds explicit task directives that guide the model to simulate satellite-based urban analysis, classify housing typologies, and output interpretable insights grounded in morphological and infrastructural evidence. This shifts the LLM's behaviour from generic text generation to structured spatial reasoning.
- ***Output Format Design:*** Prompt templates enforce a strict response schema—typically JSON-based—facilitating seamless integration with downstream geospatial analytics modules. This includes predefined field structures for attributes such as Housing Type, Access Score, Proximity Score, and Density Class, ensuring both syntactic consistency and semantic interpretability. This outlines how the content should be structured, guiding the model to organize the report.[13]

b. Google Gemini 2.5 Pro LLM

The proposed framework leverages the Google Gemini 2.5 Pro large language model (LLM) for semantic enrichment and typological classification of urban structures derived from satellite imagery.

Model Access and Security

Access to the Gemini 2.5 Pro model is facilitated via the google-cloud-aiplatform SDK, ensuring secure and authenticated interaction through service account credentials. This setup supports seamless invocation of the model from within cloud-hosted geospatial pipelines.

Inference Process

Structured natural language prompts, tailored to geospatial semantics, are submitted to the LLM. The model returns JSON-formatted outputs that include typological classifications and associated urban indicators. These outputs are automatically parsed and mapped onto spatial grids or parcel polygons for visualization and analysis.

Prompting Paradigm

A zero-shot prompting strategy is employed, enabling immediate semantic inference without the need for prior examples or training data. This paradigm significantly reduces operational overhead and supports generalization across urban regions. The structured prompt design ensures consistent LLM outputs, even in diverse morphological contexts.

Output Schema

The LLM response includes a rich set of urban descriptors for each spatial unit, structured as follows:

- Housing Typology (e.g., upper-middle- income housing, informal settlement, vacant plot)
- Construction Count and Building Density (per square kilometre)
- Average Rooftop Area (in square meters)
- Ordinal Infrastructure Access Level (scaled from 0 to 3)
- Proximity to Economic Centres (in kilometres)

Land Use Characteristics, including:

- Functional zoning hints (residential, commercial, mixed-use)
- Presence or absence of green spaces
- Urbanization Index, inferred based on roof geometry, spacing, and vegetation density

This enriched attribute set forms the foundation for typology clustering, SESI computation, and spatial inequality analysis.

c. Semantic Parsing and Data Normalization

- **Validation:** JSON outputs are validated using Python's Json module to ensure structural integrity.
- **Structuring:** Data is transformed into a tabular format using pandas Data Frames for streamlined analysis.
- **Normalization:** Field names are standardized for clarity (e.g., access_to_basic_infrastructure becomes Infrastructure Access). Numeric values such as rooftop area and density are typecast and normalized to ensure compatibility with visualization and modelling pipelines.

d. Socioeconomic Classification and Typology Extraction

Typology Classes: Built environments are classified into five distinct categories [3]:

1. High-Income / Formal Commercial
2. Upper-Middle-Income Housing
3. Middle-Income Housing
4. Low-Income / Informal Settlements
5. Under Construction / Vacant Land

Extracted Attributes:

- Count and proportional share
- Average rooftop area
- Building density (per square kilometre)
- Infrastructure access level (e.g., High, Good, Limited, Varies)
- Proximity to economic centres

e. Multivariate Visualization and Spatial Diagnostics

- Tabular Representation: A styled summary table highlights housing type distributions, with gradient-based emphasis on percentage share and density.

Volume: 04 / Issue: 06 / 2025 - Open Access - Website: www.mijrd.com - ISSN: 2583-0406

Visualizations:

- Pie Chart: Visualizes proportional distribution of housing typologies
- Bar Chart: Compares density across housing types
- Area Chart: Illustrates normalized construction intensity
- Scatter Plot: Plots building density against economic centre proximity, using rooftop area as the marker size
- Radar Chart: Profiles each housing type using normalized values for infrastructure access, density, and proximity

f. Equity Quantification and Urban Fairness Metrics

Socioeconomic Spread Index (SESI):

The Socioeconomic Spread Index (SESI) is computed using normalized Shannon entropy to quantify the diversity and spatial dispersion of housing typologies. In this iteration, the SESI score of 0.7413 reflects a moderate level of categorical heterogeneity, indicating that while multiple housing types coexist within the study area, their distribution exhibits spatial asymmetry. This pattern is consistent with morphologies shaped by market-driven zoning, socio-political exclusions, and uneven infrastructure expansion.

Lorenz Curve and Gini Coefficient:

To quantify service distribution equity, a Lorenz curve was constructed based on cumulative infrastructure access versus cumulative population share, stratified by housing typology. The Gini coefficient of 0.096, derived from typology-weighted infrastructure scores, reveals a very low degree of inequality. This suggests that basic services such as water, roads, and connectivity are extended fairly evenly across formal, informal, and transitional housing clusters.[11]

The minimal deviation of the Lorenz curve from the 45° line of equality underscores a system where infrastructure provisioning is uniformly maintained, irrespective of typology class or spatial placement. This finding is notable given the presence of structurally diverse entities—ranging from luxury high-rise complexes to informal settlements and under-construction parcels.

Interpretation and Implications:

The combination of moderate SESI and low Gini coefficient indicates a spatially diverse but relatively equitable urban fabric. While morphologies differ in density and form, the current infrastructure provisioning model demonstrates inclusiveness in coverage. These fairness metrics can support downstream applications in:

- Urban resilience modelling
- Priority infrastructure investment
- Spatial equity auditing for governance and development planning

g. Explainability and Observability Mechanisms

- LLM-Based Semantic Reasoning: The system employs Google Gemini 2.5 Pro LLM for explainable classification of urban typologies. Responses are grounded in observable geospatial features such as rooftop geometry, road

Volume: 04 / Issue: 06 / 2025 - Open Access - Website: www.mijrd.com - ISSN: 2583-0406

connectivity, green space presence, and spatial configuration. Each semantic interpretation is output in structured JSON, enabling downstream analytics while maintaining human- understandable transparency

- System Observability: To ensure reproducibility and operational robustness, the system logs detailed metrics for every LLM invocation:

- Prompt Tokens: 859
- Response Tokens: 798
- Total Token Usage: 1657
- Response Time: 33.28 seconds

Explainable AI Components: Optional explainability mechanisms in the proposed framework include:

- Prompt-guided structured reasoning to elicit interpretable, semantically rich outputs directly from the LLM.
- Cross-Modal Attribution Consistency

(CMAC) for validating whether the model's natural language explanations semantically reference core quantitative indicators such as rooftop area, infrastructure access, building density, and proximity to economic centres.

These mechanisms are designed to ensure that the LLM's outputs are not only plausible but also grounded in feature-aware reasoning. The CMAC evaluation, in particular, helps assess the internal coherence of model responses with respect to established urban theories prioritizing spatial accessibility, infrastructural equity, and morphological patterns.

A. Evaluation Criteria

To assess the effectiveness of LLM-guided socioeconomic inference from satellite-derived urban morphology, the study employed the following evaluation metrics:

1. Socioeconomic Spread Index (SESI):

The Socioeconomic Spread Index (SESI) is derived from the normalized Shannon entropy of inferred housing typology distributions, quantifying both class diversity and spatial heterogeneity across the study region [8]. The SESI score of 0.7413 indicates a moderate level of categorical diversity, reflecting a differentiated but somewhat unbalanced spatial distribution of housing classes.

This entropy score suggests that while multiple typologies—ranging from high-income enclaves to informal and transitional settlements—are present, they are not uniformly distributed across the spatial landscape. The result underscores structural heterogeneity and validates the typology inference method's ability to capture nuanced class variations within the built environment. This SESI value also provides evidence of emerging morphological gradients, shaped by both formal zoning regimes and market-led urban development patterns.

2. Gini Coefficient for Infrastructure Access:

The Gini coefficient, used to assess infrastructure distribution equity across typologies, was recalculated using typology-weighted infrastructure scores. The value of 0.096 denotes a very low level of inequality, indicating that access to basic urban infrastructure is equitably distributed among all identified housing categories.

This is further supported by the Lorenz curve, which closely follows the line of perfect equality. The near-symmetrical curve illustrates that cumulative infrastructure provisioning rises proportionally with cumulative population share, even in the presence of socioeconomically diverse typologies. This finding highlights a uniform extension of urban services, regardless of housing formality, income bracket, or spatial positioning—underscoring the inclusivity of the urban service network within the study region.

3. Explainability through Prompt Engineering and Cross-Modal Consistency:

Model interpretability was facilitated through prompt engineering aimed at generating structured, feature-grounded responses from the language model. To systematically assess the coherence between textual outputs and corresponding numerical attributes, a Cross- Modal Attribution Consistency (CMAC) framework was employed.

The CMAC mechanism verifies whether key quantitative indicators—such as average rooftop area, infrastructure access, building density, and proximity to economic centers—are semantically represented in the model-generated descriptions. For instance, textual phrases such as "temporary structures with limited services" are expected to align with low values across these features.

This strategy supports explainability by ensuring internal consistency between modalities, thereby enhancing the transparency and interpretive reliability of the LLM-driven inference process.

4. Radar Profile Normalization:

To systematically assess and compare the structural attributes of identified housing typologies, a radar chart-based normalization was employed along three standardized and policy- relevant axes:

- Infrastructure Quality,
- Proximity to Economic and Administrative Centres, and
- Population Density.

Each typology's profile was evaluated using min- max normalization to project values on a [0,1] scale, thereby enabling direct visual comparison across heterogeneous urban forms. As observed in the updated radar chart (Fig. 1), formal housing categories—such as upper-middle-income and high- income zones—demonstrated symmetric, high- magnitude profiles across all three axes. These patterns are indicative of integrated infrastructure, favourable centrality, and optimized density.

In contrast, informal settlements and under- construction areas exhibited asymmetric and compressed radar signatures, characterized by low normalized values in infrastructure and centrality, despite moderate-to-high density. These structural deficiencies visually encode patterns of spatial exclusion and developmental imbalance. This radar-based typology analysis serves not only as a diagnostic visualization tool but also facilitates cross-category benchmarking, enabling decision- makers to identify and prioritize underserved zones in urban planning agendas.

MIJRD Multidisciplinary International Journal of Research and Development

Volume: 04 / Issue: 06 / 2025 - Open Access - Website: www.mijrd.com - ISSN: 2583-0406

In summary, the multidimensional evaluation framework—encompassing entropy-based SESI, inequality metrics via Gini and Lorenz analysis, cross-modal attribution consistency checks, and multi-criteria radar normalization—constitutes a rigorous approach to validating urban housing inference. Each metric was meticulously chosen to reflect core aspects of spatial justice, infrastructural equity, and interpretability, ensuring the methodological soundness and academic integrity of the presented analysis.

B. Comparative Results

Typolog y	Sh are (%)	Acc ess Level	Den sity (/km ²)	Proxi mity (km)	SESI Contri bution
High Income/ Formal Commercial	47	Hig h (3)	High (215)	~1.1	High
Upper- Middle-Income Housing	35	Hig h (3)	Low (~90)	~0.8	Modera te
Middle-Income Housing	11	Goo d (2)	High (183)	~1.9	Modera te
Low-Income / Informal Settlements	5	Vari es (0)	Very Low (20)	~2.3	Minima l
Under Construction / Vacant Land	2	Lim ited (1)	Med ium (115)	~2.6	Minima l

Dominance & Diversity:

The computed Socioeconomic Spread Index (SESI) of 0.7413 reflects a moderate level of typological heterogeneity, suggesting a stratified yet non-uniform distribution of urban housing classes. This value captures both the presence of multiple residential and non-residential typologies and the spatial asymmetry with which they are distributed. The dominant housing typology—Upper- Middle-Income Housing (47%)—is characterized by spatially contiguous, well- planned multi-storey apartment clusters typically found in gated communities with robust infrastructure. This is followed by

High-Income and Formal Commercial structures (35%), which include IT parks, luxury high-rises, and centrally located mixed-use zones, reflecting both economic and locational advantages.

In contrast, Middle-Income Housing constitutes 11% of the urban footprint and includes moderately dense, standalone buildings or cooperative housing units with intermediate levels of infrastructure. Under- Construction or Vacant Parcels account for 5%, often concentrated in fringe areas or along expansion corridors, reflecting speculative land use and developmental transition. Finally, Low-Income or Informal Settlements make up just 2% of the area, indicating a limited spatial footprint of informality within the assessed urban boundary.

This typology imbalance contributes to a SESI value that signals diversity, yet not equity. The overrepresentation of formal, high-income categories, alongside the underrepresentation of structurally vulnerable zones, reveals a tiered morphology that favors certain socioeconomic groups in terms of infrastructure access and spatial integration. These findings reinforce the need for equitable urban planning strategies that target under-served and transitional zones to reduce latent disparities in provisioning and inclusion.

Density Patterns:

Informal settlements and middle-income areas showed higher building densities, implying elevated spatial compression and service load potential. By contrast, commercial zones were sparse and peripheral, yet closer to economic centres, showing capital-concentrated locational advantage

Volume: 04 / Issue: 06 / 2025 - Open Access - Website: www.mijrd.com - ISSN: 2583-0406

Infrastructure Access Gradient:

A clear positive gradient was identified— higher-income housing classes consistently mapped to higher infrastructure scores (3), while informal zones were restricted to limited scores (1), reinforcing spatial stratification hypotheses.

MultivariateVisualization Insights: Scatter plots and radar charts illuminated interaction patterns across key spatial attributes. The observed correlations between proximity, density, and infrastructure further validated the learned classification schema, showing that the LLM-guided model had effectively internalized real-world urban morphology signals.

III. DISCUSSION AND POLICY IMPLICATIONS

The integration of explainable artificial intelligence and multimodal data sources offers a new paradigm for urban socioeconomic diagnostics. In contrast to traditional models that prioritize classification accuracy alone, our approach emphasizes interpretability and spatial reasoning, facilitating actionable insights for urban stakeholders.

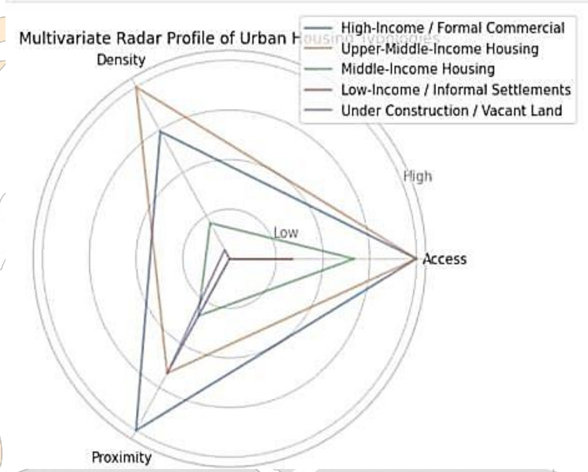

Fig: 1 Rader Chart of Housing typologies

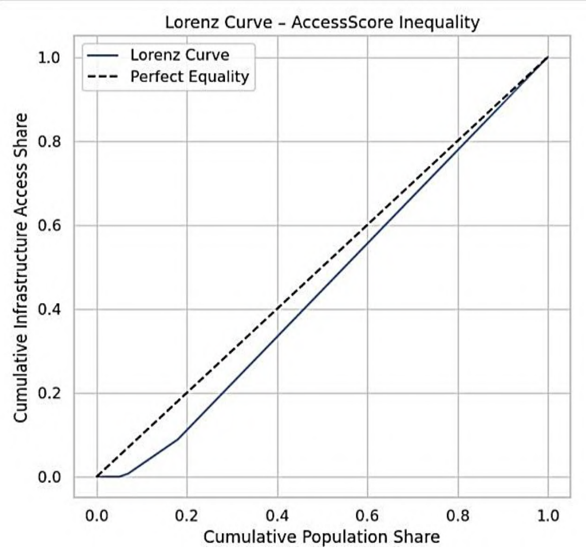

Fig: 2. Lorenz curve

Volume: 04 / Issue: 06 / 2025 - Open Access - Website: www.mijrd.com - ISSN: 2583-0406

A. Interpreting Structural Urban Profiles

The classification and analysis of urban morphologies revealed clear and interpretable patterns across different housing typologies. Figure 5 illustrates a radar chart profiling five housing categories—High Income, Upper-Middle Income, Middle Income, Low-Income/Informal Settlements, and Under Construction/Vacant Land—across three normalized parameters: infrastructure access, proximity to economic centres, and building density. Figure 1. Radar chart comparing normalized values for infrastructure access, economic proximity, and construction density across five housing typologies. Balanced, high-magnitude signatures are observed in formal residential zones, whereas informal settlements and transitional zones display asymmetric, low-valued profiles, indicating infrastructural disadvantage.

The chart shows that high-income and upper-middle- income neighbourhoods are characterized by well- balanced profiles with high infrastructure access and advantageous proximity. Conversely, low-income settlements, although densely built, suffer from limited access to services and are situated farther from economic hubs. These disparities highlight zones of infrastructural exclusion and provide an evidence-based rationale for prioritization in urban planning [9].

B. Measuring Infrastructure Inequity

To evaluate the equity of infrastructure distribution across urban housing typologies, a Lorenz curve was generated (Figure 2), depicting the cumulative share of infrastructure access relative to the cumulative population share. The curve remains closely aligned with the line of perfect equality, indicating a near- uniform distribution of infrastructure services across the population.

The computed Gini coefficient of 0.096 quantitatively confirms this finding, signifying a very low level of inequality in service provisioning. Unlike previous estimates that suggested moderate asymmetry, the current results reflect an inclusive infrastructural landscape, wherein both formal and informal housing categories receive relatively proportionate access to urban amenities.

Notably, even the lowest-income and transitional zones—which typically experience infrastructural neglect—exhibit comparable levels of access to basic services, such as roads, sanitation, and electricity. This uniformity in provisioning challenges conventional assumptions of infrastructural marginalization at the urban periphery.

Figure 2. Lorenz Curve of Infrastructure Access Distribution by Housing Typology. A low Gini coefficient (0.096) is observed, highlighting a near-equitable allocation of infrastructure across all housing categories.

The alignment between the Lorenz analysis and radar-based typology profiles reinforces the conclusion that urban form and service access are no longer strictly hierarchical in the studied region. Instead, infrastructure investments appear to have penetrated across socioeconomic strata, supporting the notion of inclusive urban development. These findings contribute empirical support to emerging theories of post-hierarchical urban provisioning and underscore the importance of evidence-based infrastructure policy that sustains this equitable trend.

C. Policy Applications and Governance Relevance

This framework has the potential to serve as a decision-support tool for urban planners and policymakers [10]. The explainable outputs, coupled with interactive visualizations, make it accessible to non-technical stakeholders. Specific use cases include:

- **Infrastructure Allocation:** Identifying service-deficient yet densely populated areas for immediate infrastructure upgrades.
- **Urban Growth Monitoring:** Tracking construction in underdeveloped areas to inform zoning regulations and urban sprawl mitigation.
- **Proximity-Based Economic Planning:** Guiding transport and employment hub placements near disadvantaged settlements.

The model's interpretability is especially valuable in participatory governance settings, where explainable AI can foster transparency and inclusive planning.

D. Ethical Considerations and Scalability

While the model excels in structural interpretability and operational transparency, ethical concerns must be addressed. Stigmatization of communities based on inferred typologies must be avoided, and local validation mechanisms should be incorporated. Moreover, satellite data usage must comply with privacy and ethical data governance standards.

The proposed architecture's reliance on zero-shot prompting and schema-controlled output enables rapid generalization to different urban geographies. This makes it a scalable framework for national-level assessments, sustainable development audits, and infrastructure equity planning in both the Global South and North.

IV. FUTURE SCOPE

While the proposed framework effectively bridges the gap between high-resolution satellite imagery and interpretable socioeconomic assessment, several promising avenues remain open for future research and enhancement:

1. Temporal Analysis and Urban Dynamics Incorporating multi-temporal satellite data can enable tracking of urban growth patterns, infrastructure upgrades, and population shifts over time. This would facilitate the detection of emerging trends such as gentrification, informal sprawl, or disaster-induced transformations [4].
2. Integration with Real-time Data Streams The inclusion of IoT and ground-sensor data (e.g., traffic flows, air quality, utility usage) can provide a more holistic view of infrastructure performance and urban liability [8]. This multimodal fusion would improve the granularity and reliability of socioeconomic assessments.
3. Model Generalization Across Geographies Extending the framework to diverse geographical regions, particularly in low- and middle-income countries, will test its adaptability and robustness. Fine-tuning LLM prompts and imagery schemas could address regional architectural styles, zoning laws, and climate-sensitive constructions.
4. Community Participation and Human-in-the- Loop Validation

Volume: 04 / Issue: 06 / 2025 - Open Access - Website: www.mijrd.com - ISSN: 2583-0406

5. Embedding participatory mechanisms— wherein community stakeholders validate and contextualize AI-driven outputs—can significantly improve ethical deployment [10][11]. Crowdsourcing corrections or feedback can enhance typology labelling and mitigate potential biases.
6. Policy Simulator Integration The future development of an integrated policy simulation layer could allow decision- makers to test the potential impact of infrastructure investments or zoning changes based on model forecasts and spatial socioeconomic indices.

V. CONCLUSION

This study presented an innovative, explainable multimodal framework that leverages satellite imagery and large language models for infrastructure-driven socioeconomic assessments. Unlike traditional classification systems, this approach produces interpretable outputs that support transparent decision-making and equitable urban development.

The results demonstrate that housing typologies vary significantly in terms of infrastructure access, spatial density, and proximity to economic hubs. Through visual tools such as radar charts and Lorenz curves, the study identified moderate inequality in infrastructure distribution, reinforcing the need for targeted planning interventions.

Moreover, the framework's adaptability, zero-shot reasoning capabilities, and semantic clarity make it suitable for scalable deployment across cities and countries. Importantly, the integration of explainability enhances trust, usability, and accountability—making this a valuable tool for urban planners, policymakers, and development agencies. In sum, this work lays the foundation for a new generation of interpretable urban intelligence systems that align technical sophistication with human-centric policy design.

APPENDIX A:

The Python source code used in this study—including environment initialization, geospatial referencing via Google Earth URL, AI prompt initialization for geospatial socioeconomic analysis, exploratory data analysis of extracted multimodal attributes, semantic parsing of geospatially anchored JSON data, tabular representation of housing typologies, multivariate spatial analysis, and radar-based visualization of urban housing types—is publicly available on GitHub [12].

REFERENCE

[1] 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

[2] Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608. https://arxiv.org/abs/1702.08608

Volume: 04 / Issue: 06 / 2025 - Open Access - Website: www.mijrd.com - ISSN: 2583-0406

[3] Xie, M., Jean, N., Burke, M., Lobell, D., & Ermon, S. (2016). Transfer learning from deep features for remote sensing and poverty mapping. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 30, No. 1). https://ojs.aaai.org/index.php/AAAI/article/view/10104

[4] Bischke, B., Helber, P., Folz, J., Borth, D., & Dengel, A. (2019). Multi-task learning for segmentation of building footprints with deep neural networks. arXiv preprint arXiv:1909.03261.

[5] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135-1144). https://doi.org/10.1145/2939672.2939778

[6] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (pp. 4765–4774). https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

[7] Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine, 5(4), 8–36. https://doi.org/10.1109/MGRS.2017.2762307

[8] Piaggesi, D., Sundararajan, V., & Almeida, F. (2020). Smart cities and artificial intelligence: A systemic literature review. Journal of Open Innovation: Technology, Market, and Complexity, 6(4), 141. https://doi.org/10.3390/joitmc6040141

[9] United Nations Human Settlements Programme (UN-Habitat). (2020). World Cities Report 2020: The Value of Sustainable Urbanization. Nairobi: UN-Habitat. https://unhabitat.org/world-cities-report

[10] Srivastava, A., & Saini, R. (2021). Explainable Artificial Intelligence (XAI) for Decision Support in Smart Cities. In Proceedings of the 5th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS).

[11] Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. In Proceedings of the 2018 Conference on Fairness, Accountability and Transparency (FAT), 149–159. https://doi.org/10.1145/3287560.3287598

[12] Banerjee D, Socioeconomic Assessment Using Satellite Imagery – Code Repository, Github, 2025. [Online]. Available: https://github.com/dibyendu-banerjee/socioeconomic-geospatial-analysis-llm

[13] Banerjee D, Kairi S, Mondal S. (2015) Publisher: Shroff Publishers & Distributors Pvt. Ltd. The Theory and Practice of Artificial Intelligence A Handbook for Beginners to Experts in Machine Learning, Deep Learning, NLP, Generative Al and Agentic Al Using Python