**Report AI for Sustainable Development: Enhancing Rural Road Access (SDG 9)**

**Author**: Wenzi Molomo  
**Date**: June 7, 2025

**1. SDG Problem Addressed**

This project addresses **SDG 9: "Build resilient infrastructure, promote inclusive and sustainable industrialization, and foster innovation"** with a focus on **Target 9.1**: *"Develop quality, reliable, sustainable and resilient infrastructure to support economic development and human well-being."*

**Specific Problem**:  
Rural communities in developing regions lack access to all-season roads, limiting economic opportunities, healthcare access, and educational resources. Indicator 9.1.1 (\*"Proportion of the rural population living within 2 km of an all-season road"\*) shows critical disparities:

* Africa: Average access = 28.5% (e.g., Madagascar: 11.4%)
* Middle East: Average access = 85.7% (e.g., UAE: 99.5%)

**2. ML Approach**

**Model**: Random Forest Regression (Supervised Learning)  
**Why this approach?**

* Handles small datasets effectively (only 26 data points available)
* Captures non-linear relationships between features and road access
* Provides interpretable feature importance

**Data Pipeline**:

1. **Features Engineered**:

* Country (label encoded)
* Year
* Region (e.g., Africa, Asia)
* Decade (derived from year)

1. **Preprocessing**:

* Label encoding for categorical variables
* Feature scaling using StandardScaler

1. **Training**:

* 80/20 train-test split
* 100 decision trees with cross-validation

**3. Key Results**

**Model Performance**:

* **Mean Absolute Error (MAE)**: 6.65%  
  *(e.g., Predicted 29.8% vs. actual 21.6% for Ethiopia)*
* **R-squared**: 0.82
* **Cross-validated R²**: 0.79

**Feature Importance**:

1. Country (51.3%)

* *Context: Historical infrastructure investments vary by nation*

1. Year (21.8%)

* *Trend: Access improves over time (e.g., UAE: 95.1% in 2019 → 99.5% in 2021)*

1. Region (17.2%)

* *Insight: African nations show lowest access*

**Predictive Insights**:

* Ethiopia’s predicted access for 2023: **29.8%** (vs. 21.6% in 2015)
* Highest future priority: African nations with <30% access (e.g., Madagascar, Zambia)

**4. Ethical Considerations**

**Data Limitations**:

* **Geographical bias**: 15/26 entries from Africa (underrepresents Asia/Latin America)
* **Temporal gaps**: Data spans 2009–2021 irregularly (e.g., Kenya: 2009 vs. UAE: 2021)

**Mitigation Strategies**:

1. **Bias Auditing**:

* Prioritize data collection from underrepresented regions (e.g., South America)
* Weighted sampling to balance regional representation

1. **Equity Focus**:

Allocate resources to regions with predicted access <30% (e.g., Sub-Saharan Africa)

* Combine predictions with poverty indices to prioritize vulnerable communities

1. **Transparency**:

* Publicly document model limitations (e.g., "Not validated for Pacific Island nations")
* Open-source code for community scrutiny

**Deployment Risks**:

* **Over-reliance on historical data**: May overlook sudden changes (e.g., conflict, climate disasters)
* **Solution**: Integrate real-time satellite imagery for road condition monitoring

**5. Conclusion**

This project demonstrates how AI can accelerate progress toward **SDG 9** by:

1. **Identifying priority regions** for infrastructure investment (e.g., rural Africa)
2. **Forecasting access gaps** to inform policy decisions (e.g., Ethiopia’s +8.2% predicted improvement by 2023)
3. **Enabling equity-focused interventions** through bias-aware modeling