**Proposal: Analyzing the Electricity Tariff Reform and Social Equity Using Machine Learning**

**Project Title**

**"Balancing Social Equity and Financial Sustainability: A Machine Learning Approach to Electricity Tariff Reforms in Sub-Saharan Africa"**

**Context and Challenge**

In Sub-Saharan Africa, electricity tariff reforms are underway to address the financial sustainability of energy providers while introducing social tariffs to protect low-income households. These reforms involve increasing overall electricity tariffs while targeting subsidies for the poorest households.

The challenge lies in evaluating the impact of these reforms on social equity and economic inequality. Specifically, this project aims to analyze how tariff changes affect household energy affordability, consumption patterns, and socio-economic disparities. By leveraging machine learning, the study seeks to identify vulnerable populations and develop data-driven recommendations for designing socially equitable tariff policies.

**Beneficiaries**

1. **Direct Beneficiaries**:
   * Low-income households in Sub-Saharan Africa, particularly in Kenya, Ethiopia, and Uganda, who are directly affected by electricity tariff reforms.
2. **Indirect Beneficiaries**:
   * Governments, energy regulators, and development organizations involved in designing and implementing energy pricing strategies.
   * International development agencies focused on promoting energy access and equity.

**Goals and Objectives**

1. **Evaluate Social Equity**: Assess how tariff reforms impact different income groups, particularly low-income households, in terms of energy affordability and consumption patterns.
2. **Identify Vulnerable Groups**: Use machine learning to identify households at risk of energy poverty due to tariff increases.
3. **Inform Policy Design**: Provide actionable insights to policymakers for balancing social equity with financial sustainability in tariff reforms.
4. **Promote Research Contributions**: Develop a robust framework for evaluating the social impact of tariff reforms, contributing to global discussions on energy pricing and equity.

**Data Details**

* **Primary Data Source**:
  + The **World Bank’s Poverty and Equity Data** (publicly available).
* **Data Features**:
  + **Income and Expenditure**: Household income and electricity expenditure as a percentage of income.
  + **Electricity Consumption**: Historical data on electricity usage at the household level.
  + **Tariff Information**: Historical and reformed tariff structures, including targeted subsidies.
  + **Demographic Data**: Household size, employment status, urban/rural classification, and geographic region.
* **Data Size**: Approximately 300,000–500,000 records, equivalent to 5–30 GB.

**Proposed Methodology**

**Phase 1: Data Collection and Preprocessing**

* **Data Extraction**: Compile data from the World Bank Poverty and Equity Data repository.
* **Cleaning and Normalization**: Handle missing values, standardize tariff categories, and normalize household income and expenditure data.
* **Feature Engineering**: Create new features, such as affordability indices, energy vulnerability scores, and socio-economic disparity metrics.

**Phase 2: Exploratory Data Analysis (EDA)**

* Use visualization and statistical techniques to identify patterns and trends in electricity consumption, affordability, and socio-economic disparities across income groups.

**Phase 3: Machine Learning Model Development**

1. **Supervised Learning Models**
   * Train models like Random Forests, XGBoost, and Gradient Boosting to predict the affordability index based on income, consumption, and demographic factors.
   * Use classification models to identify households most likely to experience energy poverty due to tariff reforms.
2. **Unsupervised Learning Models**
   * Apply clustering algorithms (e.g., K-means, Hierarchical Clustering) to group households by vulnerability and energy consumption patterns.
3. **Time-Series Analysis**
   * Analyze historical tariff and consumption data using models like ARIMA and LSTM to predict future impacts of tariff changes.
4. **Causal Inference Techniques**
   * Use Propensity Score Matching (PSM) and Double Machine Learning (DML) to estimate the causal impact of tariff reforms on household energy affordability.
5. **Explainable AI (XAI)**
   * Leverage tools such as SHAP (SHapley Additive exPlanations) to identify key factors influencing energy affordability and consumption patterns.

**Phase 4: Policy Recommendation and Reporting**

* Synthesize insights into actionable policy recommendations for energy regulators and governments.
* Develop an interactive visualization dashboard to present findings to stakeholders.

**Phase 5: Research Paper and Dissemination**

* Publish a research paper highlighting the methodology, key findings, and implications of the study.
* Present findings at energy policy and development forums to inform global discussions on social equity in energy pricing.

**Deliverables**

1. **Machine Learning Models**: Trained models for predicting energy affordability and identifying vulnerable households.
2. **Dashboard**: Interactive visualization of key insights and policy implications.
3. **Policy Recommendations**: Evidence-based strategies for designing socially equitable electricity tariff reforms.
4. **Research Paper**: Peer-reviewed publication detailing methodology, findings, and policy implications.

**Expected Outcomes**

1. **Improved Equity Analysis**: A clear understanding of how tariff reforms impact energy affordability and socio-economic disparities.
2. **Data-Driven Policy Design**: Actionable insights to support the design of targeted subsidies and equitable tariff structures.
3. **Research Contribution**: A robust framework for evaluating energy pricing policies, contributing to global discussions on energy access and equity.
4. **Enhanced Stakeholder Engagement**: A dashboard and research outputs to facilitate collaboration between governments, regulators, and development organizations.