**Advancing SDG 13: Climate Action Through Machine Learning Innovation**

Climate change is a complex global crisis, and data-driven decision-making has become a critical tool in our collective response. As nations commit to reducing greenhouse gas emissions under frameworks such as the Paris Agreement, the ability to accurately forecast CO₂ emissions is more important than ever. Our project bridges the gap between AI innovation and climate sustainability by developing a machine learning model that predicts CO₂ emissions using socio-economic data.

This work contributes directly to United Nations Sustainable Development Goal 13 (SDG 13): Climate Action, which calls for urgent action to combat climate change and its impacts.

**Project Objective: Why Predicting Emissions Matters**

CO₂ is the principal greenhouse gas contributing to global warming. Understanding its sources and trends enables proactive policymaking and targeted intervention. However, traditional forecasting approaches are often limited in their ability to handle non-linear relationships, high-dimensional data, or interacting factors.

Our project aims to overcome these challenges by leveraging machine learning, a subset of artificial intelligence capable of learning complex patterns from data.

Specifically, this project seeks to:

* Develop a predictive model for CO₂ emissions based on macroeconomic and energy indicators
* Identify key features (factors) that drive carbon emissions
* Provide a tool for evidence-based environmental policy design
* Enable what-if analyses for assessing the impact of regulatory scenarios

**Technical Approach: Machine Learning in Action**

Our project uses a supervised machine learning algorithm—Random Forest Regression—to forecast CO₂ emissions. Supervised learning involves training a model on labeled data, where the input features and the target output (CO₂ emissions) are known. Once trained, the model can generalize its learning to make predictions on new, unseen data.

✳️ Why Random Forest?

Random Forest is an ensemble method that builds multiple decision trees and merges them to improve prediction accuracy and control overfitting. It is particularly suitable for problems with complex, non-linear relationships — such as the interaction between economic activity and environmental impact.

**🧩 Features Used (Input Variables):**

* Gross Domestic Product (GDP) – A measure of economic output
* Population Size – Influences overall consumption and emissions
* Energy Consumption – Drives industrial and residential CO₂ emissions
* Industrial Production Index – Proxy for manufacturing-related emissions
* Renewable Energy Share (%) – Indicator of sustainable energy adoption

📐 Evaluation Metrics:

To assess model performance, we used:

* R² (Coefficient of Determination): Indicates how well predictions approximate actual outcomes
* MAE (Mean Absolute Error): Measures the average magnitude of prediction errors
* RMSE (Root Mean Squared Error): Penalizes larger errors more heavily
* Cross-validation: Ensures the model performs consistently across data splits

**Model Performance**

The model performed strongly on the test dataset, showing both accuracy and reliability:

| Metric | Value | Change |
| --- | --- | --- |
| R² Score | 0.87 | ↑ 5% |
| MAE | 0.42 | ↓ 12% |
| RMSE | 0.58 | ↓ 8% |

These results suggest that the model effectively captures the underlying trends in CO₂ emissions based on the selected indicators. The feature importance analysis revealed energy consumption and GDP as the strongest predictors—consistent with established environmental-economic theory.

🌍 Social and Climate Impact

The practical implications of this model extend beyond the technical realm. Here's how it supports SDG 13: Climate Action:

📌 Evidence-Based Policy Design

By identifying the main drivers of emissions, policymakers can implement targeted regulations, such as taxing high-energy industries or subsidizing renewable alternatives.

🧭 Strategic Climate Planning

With forecasting capabilities, governments can simulate the impact of different policy interventions—from emissions caps to clean energy incentives—allowing more strategic climate planning.

🏙️ Community-Level Action

Regions identified as high-risk can receive priority support for climate adaptation, protecting vulnerable communities from the most severe impacts of global warming.

🔍 Transparency & Accountability

Publicly accessible prediction tools enhance transparency and build trust among citizens, NGOs, and international bodies monitoring national commitments.

**Ethical and Practical Considerations**

As with any AI solution, ethical considerations are paramount. This includes:

* Avoiding algorithmic bias by ensuring diverse and representative training data
* Protecting sensitive data and adhering to data governance policies
* Ensuring model interpretability, so stakeholders can understand and trust the predictions
* Avoiding overreliance on predictions without complementary human judgment

Machine learning should complement, not replace, the nuanced decision-making required in climate governance.

**💡 Conclusion**

This project demonstrates that machine learning is a powerful ally in the fight against climate change. By delivering actionable insights, supporting smarter policies, and enabling early intervention, this model exemplifies how AI can accelerate progress toward sustainable, data-informed climate action.