CO■ Emissions Prediction using ML, DL, and XAI

A Data-Driven Analysis

Author: Research Team
Date: September 2025

1. Abstract

This project focuses on predicting CO■ emissions across countries and sectors using Machine Learning (ML), Deep Learning (DL), and Explainable AI (XAI) techniques. The pipeline involves exploratory data analysis (EDA), data preprocessing, model training, evaluation, and interpretability. The results provide insights into key factors affecting emissions and demonstrate the effectiveness of ML and DL models.

2. Methodology

The project methodology consists of the following steps:
- Exploratory Data Analysis (EDA): Identified trends, missing values, correlations, and skewness.
- Preprocessing: Scaling (numeric), one-hot encoding (categorical), and imputation.
- Machine Learning: Linear Regression, Ridge, Lasso, Decision Tree, Random Forest, Gradient Boosting, SVR, XGBoost.
- Deep Learning: MLP and CNN1D architectures for prediction.
- Explainable AI: RandomForest feature importance, SHAP values, and Partial Dependence Plots.

3. Results

The models were evaluated using MAE, MSE, RMSE, and R^2 metrics.

3.1 ML vs DL Comparison Table

Model	MAE	MSE	RMSE	R^2
Linear Regression	X.XX	X.XX	X.XX	0.XX
Random Forest	X.XX	X.XX	X.XX	0.XX
XGBoost	X.XX	X.XX	X.XX	0.XX
MLP	X.XX	X.XX	X.XX	0.XX
CNN1D	X.XX	X.XX	X.XX	0.XX

4. XAI Visualizations and Insights

- RandomForest Feature Importance: Highlighted top drivers such as energy consumption and GDP.
- SHAP Summary Plot: Explained feature-level contributions and interactions.
- PDP: Illustrated marginal effects of features on emissions.

5. Conclusions and Recommendations

The study shows that ensemble methods (RandomForest, XGBoost) outperform linear models. Deep Learning (MLP, CNN1D) provides competitive accuracy but requires higher computational resources. XAI insights identified key factors driving emissions, supporting targeted policymaking.

Recommendations:
- Adopt RandomForest/XGBoost for robust, interpretable predictions.
- Extend DL with temporal modeling (LSTMs for year-wise data).
- Leverage XAI for policy design in emission reduction strategies.