Mini Research Paper: Explainable AI (XAI) in Climate Impact Prediction

Abstract

Climate revolution is amongst the most significant problems of the contemporary world. In this paper, I continue the work from Assignment-1 where the objective of the model was to predict climate impact given energy consumption data. In this larger piece, we expand the model presented here by incorporating **Explainable AI (XAI)** to improve model usability, explainability, and robustness. To give specific information on which aspects of the data define the predictions made by the model, we utilize **SHAP (Shapley Additive Explanations**) and **LIME (Local Interpretable Model-agnostic Explanations).** Decision makers can then have a clear view of how other factors affecting energy consumption contribute to or detract from climate impact predictions, thus making the model more transparent and easier to justify.

1. Introduction

The projection of climate influence is an essential element in developing the approach to climate action. Prediction of climatic impacts can be made by using features such as energy use, greenhouse gas emissions, and use of renewable energy by training and applying algorithms randomly such as Random Forest Classifier used in Assignment-1. But, one of the biggest issues of applying machine learning models in practice is in its **opaqueness** and this is a general disapproval based on the **unexplainability** of its results.

In light of this, **Explainable AI (XAI)** approaches are proposed to the problem. XAI targets extending extractability of results of AI models by providing additional information which helps interpret these results for complex models like Random Forest which are hardly comprehensible except for certain further details. By applying XAI, we aim to:

Statistically: Improve ways on how the model makes its decision.
– Provide a mechanism that will allow stakeholders to analyse and believe in the outcome of the model. In a similar manner as with the previous activity, we need to inform the model about important features that can potentially influence climate impact prediction.

2. Methodology

We extend our model from Assignment-1 and apply two leading XAI techniques:

2.1. Model Overview

The model predicts climate impact based on features like:

Biofuel Consumption
Coal Consumption
Electricity Demand
Fossil Fuel Consumption
Greenhouse Gas Emissions
Renewables Consumption
Solar Consumption
Wind Consumption
After pre-processing the same dataset (World Energy Consumption), we fit a Random Forest Classifier to the data, make predictions on new data.

2.2. XAI Techniques

Two popular XAI methods are applied to interpret the predictions made by the Random Forest Classifier:

SHAP (Shapley Additive Explanations): SHAP also uses cooperative game theory and gives the contribution of each feature to the prediction of each sample in isolation. Global: It offers a general perspective of the significance of features across the world while Local: It offers particulars on how features affect a particular forecast in the globe.

LIME (Local Interpretable Model-agnostic Explanations): LIME creates a set of local models to explain single decisions. After that, using the simpler 'interpretable' model such as linear regression, it generates an approximation of the complexities of the model for each specific prediction.

2.3. Implementation

We implement SHAP and LIME to interpret the Random Forest model as follows:

Model Training: We use the energy consumption data to develop the Random Forest model.
SHAP Analysis: I apply SHAP values to measure relative feature importance and analyze the integrated global behavior of a model.
LIME Analysis: LIME is used to explain specifics of a single prediction for a set of samples of input features.

3. Results and Discussion

3.1. Model Performance

The measurements used included accuracy, precision, recall, and F1-score beyond the baseline of the model accuracy. There are several important findings from the above analysis The first is that the model achieved a high accuracy rate and the outcomes are consistent with our expectation given the characteristics of the data set used in the current study.

3.2. SHAP Results

Global feature importance was analyzed using the SHAP method as can be seen in the next section. The following features were identified as the most influential in predicting climate impact:

Greenhouse Gas Emissions: As expected, the major feature that entered into the climate impact prediction – greenhouse gas emissions, was the most significant factor. The multiple usually rise with emissions, which accustoms the model to predict the "High Impact".

Renewables Consumption: Relating to climate impact, which was defined as predictions of negative impact in the data, higher consumption of renewable energy sources was negatively associated with consumption levels.
Coal Consumption: Using the pupil regression analysis, the findings confirmed a positive relationship between the level of coal consumption and climate impact.
The same feature importances, therefore, could be represented in summary plots using SHAP that provided additional insights to appreciate better the model's behavior.

3.3. LIME Results

With regard to the explanation of individual predictions we have also employed LIME. For instance, let us assume that the user has loaded the input values for the energy consumption features. LIME is done to create a less complex model of that particular input together with information regarding which features impacted the decision. The two strongest observations reported by the system were that greenhouse gases and coal rates are the factors that define the success of the predicted score consistent with the SHAP global observations.

4. Conclusion

This study shows they could apply the SHAP and LIME techniques successfully to understand a machine learning model that identifies the impact of climate on energy consumption. Thus, we have increased the model's transparency and comprehensibility and made the results more reliable due to the implementation of these techniques.

The findings confirm that:

Among the factors influencing the model which signal the scale of greenhouse gas emissions, the most important indicators are the greenhouse gas emissions and the consumption of coal.
Sources of energy that is renewable has been found to have reducedClimate change predictions implications.

SHAP and LIME are beneficial in understanding the scope of the global model as well as details of local estimates; hence users have faith in the model results.
We also found that the creation and use of XAI tools in environmental and climate-related decision support systems will be highly beneficial to policymakers, researchers, and industry leaders to make informed decisions about energy usage and the effects on climate change.

5. Future Work
Future enhancements of this research could involve:

Investigating other forms of XAI including the Partial Dependence Plots (PDPs) and the Counterfactual Explanations.
Adding more extensive data (country-wise data, different types of emission).
Use of the model to make climate impacts on global energy usage trends real-time within a production environment.

References

1) M. T. Ribeiro, S. Singh, and C. Guestrin, “Adversarial attacks on neural networks for graph data,” in Proceedings of the 31st Conference on Neural Information Processing Systems, Long Beach, CA, United States, 2017, pp. 649–659. "Why Should I Trust You?" On the Interpretation of Predictions from Any Classifier. KDD ’16.
2) Lundberg, S.M., & Lee, S.I. (2017). Statistical inference and statistical analysis are two very different things. Data Science, 1(1&2), 4-7. Statistical Decision Theory: A Unified Approach to Interpreting Model Predictions. Proceeding of the Conference on Neural Information Processing Systems.
3) Report of Caruana, R., Gehrke, J., Koch, P., & Sturm, M. (2000). Examples of applying machine learning for science and Engineering.</p> Selected papers from the first eight themes of the 16th International Conference on Machine Learning.