Shapley Regression Results

Variable	Direction	Shapley Share	p
Yield curve slope	-1	0.11427699	0.002
CPI	-1	0.07623505	0.001
Broad money	-1	0.03409193	1.000
Stock market	-1	0.01830194	0.161
Consumption	-1	0.04076334	0.000
Public debt	1	0.04625652	0.495
Investment	1	0.04339621	0.025
Current account	-1	0.04298019	0.094
Credit	1	0.10891914	0.006
Debt service ratio	1	0.06319236	0.051
Global credit	1	0.16916482	0.000
Global yield curve slope	-1	0.24242152	0.000

Table 1: Shapley regression and direction of alignment between predictor and crisis outcome

Shapley Values Formula

The Shapley value (ϕ_i) for feature i is calculated using the following formula:

$$\phi_i(f) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|! \cdot (|N| - |S| - 1)!}{|N|!} \left[f(S \cup \{i\}) - f(S)\right]$$

Here:

- N is the set of all features.
- S is a coalition of features that does not contain feature i.
- $f(S)$ is the model's prediction when considering only the features in set S.
- $f(S \cup \{i\})$ is the model's prediction when adding feature i to set S.
- $|S|$ is the number of features in set S.

Explanation of Shapley Values

Shapley values, rooted in cooperative game theory, have gained prominence in machine learning for their ability to attribute the output of a model to its individual features in a fair and consistent manner. This interpretability tool helps address the "credit assignment problem" by quantifying the contribution of each feature to the model's predictions. The Shapley value for a particular feature is computed as the average marginal contribution across all possible feature coalitions.

The practical utility of Shapley values lies in their ability to provide a fair and theoretically grounded method for feature attribution. Unlike some other attribution methods that can suffer from issues like inconsistency or lack of fairness, Shapley values distribute the credit for a model's prediction among its features in a way that is both fair and mathematically sound.

Moreover, Shapley values are model-agnostic, making them applicable to a wide range of machine learning models, including linear models, tree-based models, and complex models like neural networks. This versatility enhances their usefulness in various domains, from finance and healthcare to natural language processing.

By employing Shapley values, practitioners and researchers can gain deeper insights into the inner workings of their models, understand the relative importance of different features, and improve model transparency. This interpretability is crucial, especially in sensitive applications where decisions impact individuals' lives, as it allows stakeholders to comprehend and trust the model's predictions, fostering accountability and ethical machine learning practices.

Shapley Regression

Shapley Regression, also known as Shapley Additive Regression, extends the concept of Shapley values to the context of regression analysis. It aims to provide a fair allocation of the prediction output to each feature in a regression model, facilitating the understanding of individual feature contributions to the predicted outcome.

The Shapley Regression value (ϕ_i) for a feature i is calculated based on the Shapley values concept. Given a regression model f and a dataset with input features X and output y, the Shapley Regression value is defined as:

$$\phi_i(f) = \frac{1}{N!} \sum_{\pi \in \Pi_N} \left[f(x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(N)}) - f(x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(i-1)}, x_i, x_{\pi(i+1)}, ..., x_{\pi(N)}) \right]$$

Here:

- N is the number of features.
- $x_{\pi(j)}$ represents the j-th feature in the permutation π.
- Π_N is the set of all permutations of the features.
- $f(x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(N)})$ is the model's prediction for the permutation π.

- $f(x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(i-1)}, x_i, x_{\pi(i+1)}, ..., x_{\pi(N)})$ is the model's prediction for the permutation π with the i-th feature replaced.

In simpler terms, Shapley Regression calculates the average contribution of each feature across all possible orders of feature inclusion in the prediction. It considers every possible combination of features and calculates the marginal contribution of adding a feature to a particular permutation.

Shapley Regression is valuable in regression analysis for several reasons:

1. **Fairness and Consistency:** It provides a fair and consistent way to attribute the model's prediction to individual features.

2. **Interpretability:** It helps interpret the impact of each feature on the regression output, aiding in understanding the model's behavior.

3. **Feature Importance:** It quantifies the importance of each feature in the context of the regression model.

Implementations of Shapley Regression are available in various machine learning libraries, and they can be applied to a range of regression models, including linear regression, decision trees, and more complex models. Using Shapley Regression can enhance the interpretability of regression models and support better decision-making based on a deeper understanding of feature contributions.