

regression_cv.py

Summary

This code defines a `RegressionCV` class for automated regression tree modeling using Bayesian optimization.

Dependencies

Standard Library

- `typing`

Other

- `numpy`
- `pandas`
- `pydantic`
- `shap`
- `sklearn`
- `xgboost`

Description

The `regression_cv.py` file implements a `RegressionCV` class, which is an automated regression tree model based on Bayesian optimization. This class combines several machine learning techniques to create a powerful and flexible regression model.

The `RegressionCV` class inherits from `BaseAutoCV`, `RegressorMixin`, and `ExplainerMixin`, providing a comprehensive set of functionalities for regression tasks, model optimization, and result explanation. It uses XGBoost as the underlying regression algorithm and incorporates Shapley Additive Explanations (SHAP) for model interpretability.

The class allows for customization through a `RegressionCVConfig` dataclass, which includes options for monotonicity constraints, feature interactions, hyperparameter search spaces, and parallelization. The model optimization process uses Bayesian optimization to find the best hyperparameters within the specified search space.

Key features of the `RegressionCV` class include: 1. Automated hyperparameter tuning using Bayesian optimization 2. Support for custom evaluation metrics 3. Cross-validation during model selection 4. Feature importance calculation 5. Model explanation using SHAP values

The class provides methods for fitting the model, making predictions, and explaining the model's decisions. It also includes utility methods for input validation and scoring.

This documentation was generated using claude-3-5-sonnet-20240620