

Feature Importance and Vulnerability Analysis in ML Models

Contributor [Zaher Khateeb](#) [AI/ML Engineer](#) [LinkedIn](#) [zahere](#)

Project Overview

This repository provides a robust framework for conducting feature importance and vulnerability analysis in machine learning models, specifically designed for tabular data. The framework addresses two primary business problems:

1. **Improving the prediction of marketing campaign success for term deposits** using the [bank marketing dataset](#)
2. **Enhancing the assessment of credit risk** using the [German credit risk dataset](#)

Objectives

- Implement a robust framework for analyzing and improving machine learning models.
- Conduct uncertainty, feature importance, and feature performance analyses to identify model weaknesses and vulnerabilities.
- Enhance the performance of marketing and credit risk models by improving their feature selection and understanding the impact of individual features on model predictions.

Baselines and Metrics

1. Marketing Model:

- [Marketing Baseline](#)
- Provides insights and predictions to optimize campaign strategies, targeting the most likely customers for term deposits.
- Baseline ROC-AUC: 0.89
- Our improvement: Increase ROC-AUC by 2% (0.91)

2. Credit Risk Model:

- [Credit Risk Baseline](#)
- Offers more accurate risk assessments to inform lending decisions and reduce default rates.
- Baseline ROC-AUC: 0.78
- Our improvement: Increase ROC-AUC by 2% (0.80)

Analyses Execution

1. Iterate over the [analyses_to_run](#) list.
2. Check if each analysis is defined in the [analysis_methods](#) dictionary.
3. For each valid analysis, iterate over the [trained_pipelines](#).
4. Execute the specified analysis method on each pipeline.

Analyses Methods

Three specific types of analyses are defined in the `ModelImprover` class:

1. Uncertainty Analysis

- **Purpose:** Understand the confidence level of the model's predictions.
- **Method:** Uses a baseline ensemble Monte Carlo method to calculate the uncertainty of the model's predictions.

2. Feature Importance Analysis

- **Purpose:** Determine the importance of different features used by the model.
- **Method:** Uses SHAP values to plot feature importance and SHAP summary plots, and selects features based on their SHAP values.

3. Feature Performance Analysis

- **Purpose:** Analyze the performance of individual features in contributing to the model's predictions.
- **Method:** Assesses how changes in feature values affect model accuracy or other performance metrics, identifying weaknesses in the model's use of certain features.

Utility Classes and Methods

1. Uncertainty

- **Methods:**
 - `baseline_ensemble_monte_carlo`: Calculates the uncertainty of the model's predictions using ensemble Monte Carlo simulations.

2. Explainability

- **Methods:**
 - `plot_feature_importance`: Plots the importance of each feature.
 - `plot_shap_summary`: Creates a SHAP summary plot.
 - `select_features_based_on_shap`: Selects features based on their SHAP values.

3. FeaturePerformanceWeaknessAnalyzer

- **Methods:**
 - `analyze_feature_performance`: Analyzes the performance of individual features.
 - `plot_metric_drops`: Plots the performance drops for vulnerable features.

Getting Started

Prerequisites

Python	3.7 or higher	scikit-learn	Required	XGBoost	Required	SHAP	Required	matplotlib	Required
numpy	Required	pandas	Required	scipy	Required	joblib	Required		

Installation

1. Clone the repository:

```
git clone https://github.com/zahere/MLOps-24.git  
cd MLOps-24
```

2. Install the dependencies:

```
pip install -r requirements.txt
```

Usage

1. Prepare your datasets and ensure you add the configuration json to /config folder .
2. Define your pipelines and train your models.
3. Configure the `analyses_to_run` list and the `analysis_methods` dictionary.
4. Execute the analyses using the `ModelImprover` class.

License

This project is licensed under the MIT License - see the [LICENSE](#) file for details.

Acknowledgments

- Special thanks to Dr. Ishai Rosenberg - MLOps 24 Course (Y-Data)