Explainable AI

Abrhame Habtom

August 2025

1 Abstract

This paper emphasizes the explainability of Random Forest and Logistic Regression using SHAP.

2 Introduction

SHAP is used to explain which features of the data have a greater effect on it. Using two plot types, the force plot and waterfall plots, explains more of the model.

3 Methodology

I selected the Iris dataset, a classic tabular dataset for multiclass classification. It consists of 150 samples from three species of iris flowers (setosa, versicolor, virginica), with four features: sepal length, sepal width, petal length, and petal width (all in cm). I used two models, Logistic Regression (a linear model) and Random Forest Classifier (a tree-based ensemble model) and trained them both on the Iris dataset using SkLearn. Both models were trained on 80 percent of the data (120 samples) and evaluated on the remaining 20 percent (30 samples). Both achieved perfect accuracy (1.0) on the test set. To interpret the models, I applied SHAP (SHapley Additive Explanations) to compute feature contributions for predictions on the test set. SHAP provides local explanations that can be aggregated to understand the importance of global features (via mean absolute SHAP values) and the behavior of the model.11

4 Results

4.1 Feature Importance Comparison

Table 1: Feature Importance Comparison (Mean Absolute SHAP Values)

Feature	Logistic Regression SHAP	Random Forest SHAP
sepal length (cm)	0.238774	0.035753
sepal width (cm)	0.190218	0.007791
petal length (cm)	2.853299	0.199841
petal width (cm)	0.842775	0.197118

4.2 visualization

4.2.1 Logistic Regression

Figure 1: SHAP Force Plot for Logistic Regression

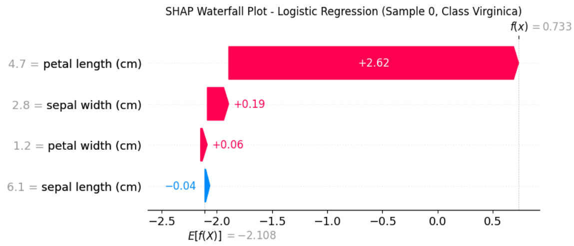

Figure 2: SHAP Waterfall Plot for Logistic Regression

4.2.2 Random Forest

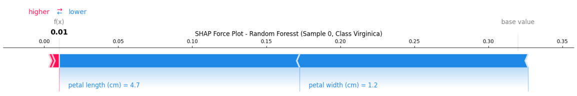

Figure 3: SHAP Force Plot for Random Forest

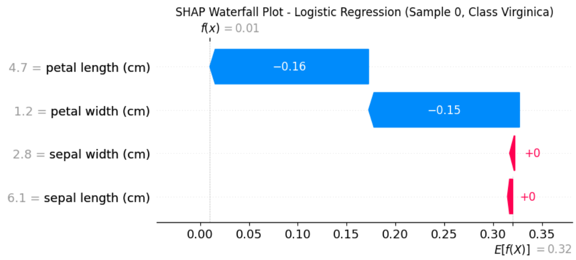

Figure 4: SHAP Waterfall Plot for Random Forest

5 Discussion

The visualization result are done on class veronica. The logistic regression is more of influenced by the three features petal length , sepal width and sepal length respectively. But the random forest result is more negatively influenced by these features. I have two reasons for the effect of the models on the feature importance

a) Feature Interaction Handling

Random Forest Can model non-linear interactions. For example, it can say "if petal length ¿ X and petal width ¿ Y, then it's Virginica". This makes the contribution of each feature more distributed.

Logistic Regression Only fits a linear hyperplane, so features that provide strong linear separation (like petal length) get disproportionately high weights.

b) Feature Scaling and Decision Paths

Logistic Regression coefficients scale with feature strength and direction, directly amplifying their SHAP effects.

Random Forest splits data across many trees, so the effect of a single feature at a single split may be diluted when averaged across many trees