



Unit-2: Explainability for Tabular Data

Permutation Feature Importance

Permutation importance for a single feature measures the decrease in a model's score when that feature's values are randomly shuffled.

When we shuffle a feature, we randomly permute its values across all samples, while keeping the other features unchanged.

Example??



Unit-2: Explainability for Tabular Data

Permutation Feature Importance

By shuffling the values of a given feature, one destroys the model's ability to make meaningful predictions using that feature.

Justify with an Example??

If the change in validation error is negligible, the feature isn't important or useful for model predictions.

Justify with an Example??

Determining the optimal number of repeats for feature permutation may require some experimentation.

If we repeat too few times, the results may be noisy and unreliable, and If we repeat too many times, it increases the computation time unnecessarily.

The correct number of repeats depends on: Randomness in data and model, Model complexity, and Available computational resources



Unit-2: Explainability for Tabular Data

Permutation Feature Importance

Permutation feature importance is a perturbation-based feature attribution technique commonly used for tabular datasets

“Perturbation” means making small, controlled changes to the input data and observing how the model’s output changes.

Each column = feature, each row = observation.

What are the efficiency factors in this method?

Model Agnostic, No retraining requirements, Simple implementation, High Interpretability, Parallelization possibility, Useful for tabular data, etc.



Unit-2: Explainability for Tabular Data

Permutation Feature Importance - Case Study Classroom Exercise

Understanding Permutation Feature Importance for Predicting House Prices

- Load a tabular dataset (California Housing dataset from scikit-learn).
- Train a regression model (Random Forest).
- Compute baseline performance (R^2).
- Shuffle each feature one by one and compute the performance drop.
- Rank features by importance.

Use `permutation_importance` from Scikit Learn



Unit-2: Explainability for Tabular Data

Permutation Feature Importance

Permutation feature importance is a perturbation-based feature attribution technique commonly used for tabular datasets

“Perturbation” means making small, controlled changes to the input data and observing how the model’s output changes.

Each column = feature, each row = observation.

What are the efficiency factors in this method?

Model Agnostic, No retraining requirements, Simple implementation, High Interpretability, Parallelization possibility, Useful for tabular data, etc.



Unit-2: SHAP (SHapley Additive exPlanations) Values

SHAP values are a specific implementation of the Shapley value calculation.

Axioms of Fairness

1. Efficiency
2. Symmetry
3. Null Player
4. Additive/Linearity

If a feature consistently has no effect on the model's output regardless of other features, its importance (Shapley value) is zero.

Example??



Unit-2: SHAP – Waterfall Visualization

Shows how each feature pushes a model's prediction from an expected baseline to the final output for a single prediction

A waterfall chart is a graphical tool primarily used to show the collective influence of successive positive and negative variables on an initial starting point.

A waterfall chart identifies trends, areas for improvement, and success factors.

XXXXX

YYYYY