

# SHAP (Shapely Additive exPlanations)

- SHAP is a unified approach to explain the output of any machine learning model
- Based on game theory and provides a way to fairly distribute the "contribution" each feature makes to a prediction
- Calculates how each feature's contribution changes across all possible combinations of other features

## SHAP Visualization Methods

- **Force Plots:** Show how each feature pushes the prediction higher or lower
- **Summary Plots:** Display overall feature importance across multiple predictions
- **Dependence Plots:** Reveal how feature interactions affect predictions
- **Decision Plots:** Trace how predictions change as features are introduced

## Why SHAP is Useful

- Model Interpretability – helps understand feature importance better
- Educational Value: provides visual explanations
- Trustworthy AI Development – Can identify potential bias in models
- Debugging Models – improves feature engineering and guides model refinement

