Model Interpretation Methods
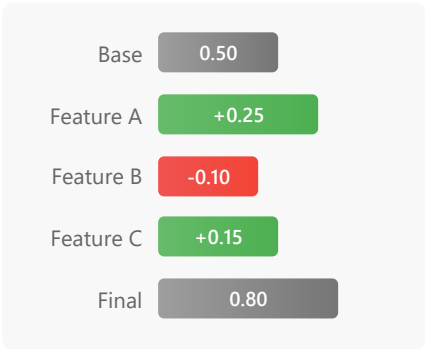
SHAP, LIME & Attention Mechanisms

SHAP Summary Plot

Age	
Income	
Education	
Experience	
Location	
Gender	
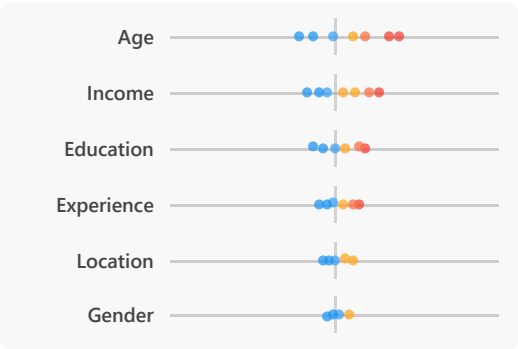
Feature importance ranking

SHAP Waterfall

Base	0.50
Feature A	+0.25
Feature B	-0.10
Feature C	+0.15
Final	0.80

Individual contribution

🔍 Methods

SHAP
- Game-theory based
- Global + local

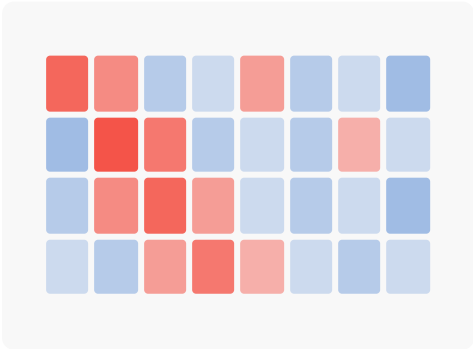**LIME**
- Model-agnostic
- Local perturbation

Attention
- Built-in mechanism
- Text/image focus

LIME Explanation

This **excellent** movie was **amazing** and **boring** scenes were few. **Great** acting!

Local linear approximation

Attention Heatmap

Model focus visualization

Scope

SHAP	Global/Local
LIME	Local only
Attention	Instance-level

✓ Best Practices

✓ Use SHAP for importance
✓ LIME for quick insights
✓ Attention for NLP/vision
✓ Compare with PDP