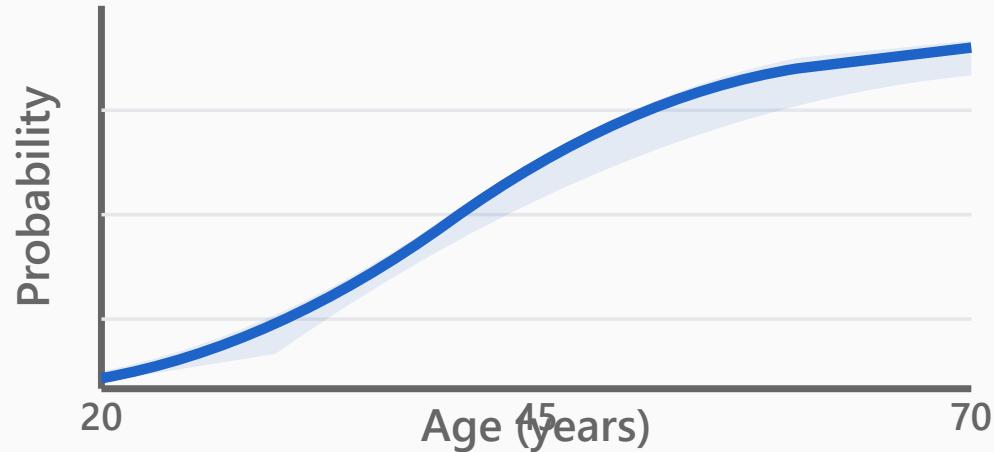


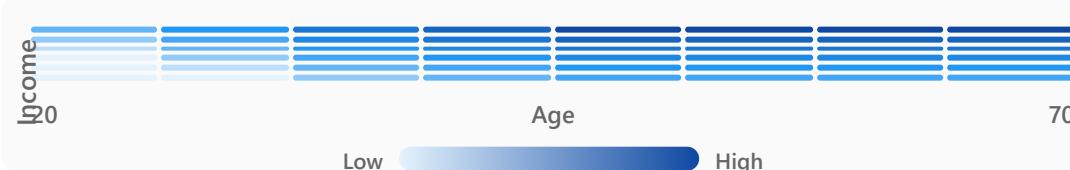
Partial Dependence Plots (PDP)

Marginal Effect Visualization

1D PDP: Age vs Loan Approval Probability



2D PDP: Age × Income Interaction



Key Concepts

Marginal Effect

Average prediction as feature varies

1D PDP

Single feature effect (line plot)

2D PDP

Two-feature interactions (heatmap)

Implementation

`sklearn.inspection.PartialDependenceDisplay`

⚠ Assumption

Features are independent - misleading if violated

Reveals

- Non-linear relationships
- Thresholds & transitions
- Average behavior patterns

How PDP is Calculated

Step-by-Step Mathematical Process

$$PDP_{xs}(x_s) = E_{X_c}[f(x_s, X_c)] = (1/n) \sum_{i=1}^n f(x_s, x_c^{(i)})$$

x_s : Target feature value (fixed) | X_c : All other features (vary across data) | f : Model prediction function | n : Number of data points

1

Original Data

Age	Inc	\hat{y}
25	40K	0.3
35	60K	0.7
45	80K	0.9
55	50K	0.6

Start with training dataset containing all features and their predictions

2

Fix Target Feature

Age	Inc	\hat{y}
30	40K	?
30	60K	?
30	80K	?
30	50K	?

Replace target feature (Age) with specific value (e.g., 30) for ALL samples

3

Predict & Average

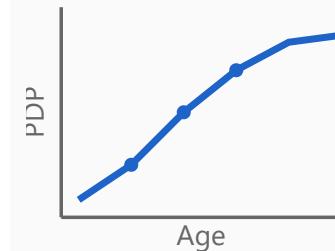
$$\begin{aligned} f(30, 40K) &= 0.45 \\ f(30, 60K) &= 0.65 \\ f(30, 80K) &= 0.75 \\ f(30, 50K) &= 0.55 \end{aligned}$$

$$\begin{aligned} PDP(\text{Age}=30) &= \\ (0.45+0.65+0.75+0.55)/4 &= 0.60 \end{aligned}$$

Get predictions for all modified samples, then compute their average

4

Repeat & Plot



Repeat steps 2-3 for different Age values to create the complete PDP curve

Key Insight

PDP marginalizes out other features by averaging predictions across their distribution. This shows the **average effect** of the target feature, assuming feature independence.