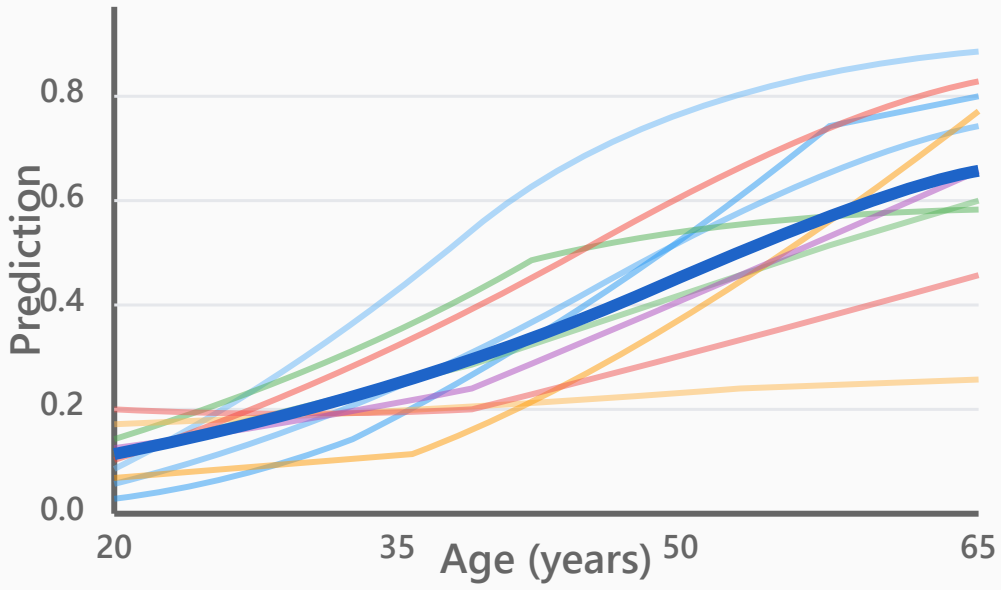
Individual Conditional Expectation (ICE)

Instance-Level Feature Effects

ICE Curves: Age vs Loan Approval Prediction



Key Concepts

ICE vs PDP

ICE
Individual instance trajectories

PDP
Average of all ICE curves

Heterogeneity

Reveals variation in feature effects across instances

Subgroups

Identifies different feature-outcome relationships

Interactions

Detects interactions PDP may hide

Non-parallel Lines

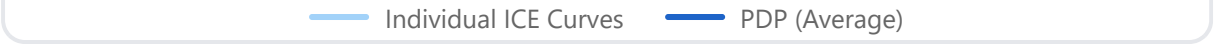
Shows unexpected patterns & heterogeneity

Centered ICE

Subtract baseline for better comparison

What to Look For

- Parallel lines = homogeneous effect
- Diverging lines = heterogeneous effect



- Clusters = distinct subgroups

How ICE Works: Step-by-Step Process

1

Select Instance

Choose a single data point from your dataset

x_1 = (age=30, income=50k, ...)

→

2

Vary Feature

Change target feature across its range, keep others fixed

age: 20, 25, 30, ..., 65
income=50k (fixed)

→

3

Get Predictions

Run model for each feature value

$\hat{y}(20)$, $\hat{y}(25)$, $\hat{y}(30)$, ..., $\hat{y}(65)$

→

4

Plot & Repeat

Draw curve for this instance, repeat for all instances

n curves for n instances
Average = PDP