Instance-Based vs. Model-Based Learning

Machine Learning Notes

1 Core Concepts

Machine learning approaches differ in how they **learn from data**:

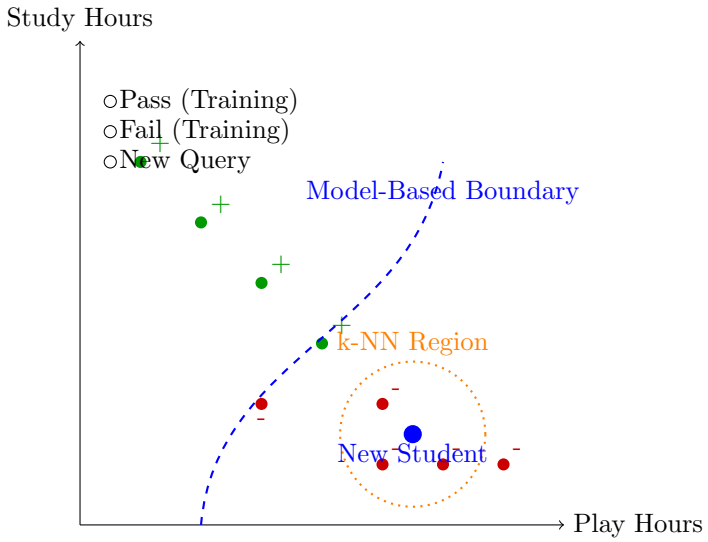
- **Instance-Based Learning**: Memorizes training data, makes predictions using direct data comparisons
- **Model-Based Learning**: Discovers underlying patterns, creates generalized decision rules

2 Key Differences

Aspect	Instance-Based	Model-Based
Learning Approach	Memorizes raw data	Learns data patterns
Training Phase	Minimal (stores data)	Parameter estimation
Storage	Entire training data	Compact model (e.g., pickle file)
Prediction Speed	Slow (data search)	Fast (formula execution)
Generalization	Before prediction	During training

3 Visual Example: Pass/Fail Prediction

Predict student success based on study/play hours:



How They Predict

- **Instance-Based (k-NN)**:

1. Find k nearest neighbors
2. Majority vote: Pass if $\sum \text{Pass}_{neighbors} > \sum \text{Fail}_{neighbors}$

- **Model-Based**:

  Use decision boundary:
  Pass if $\beta_0 + \beta_1 x_{\text{play}} + \beta_2 x_{\text{study}} > 0$

4 Technical Comparison

Workflow Differences

- **Model-Based**:
  1. Train model → discover patterns
  2. Store parameters (e.g., weights β_i)
  3. Discard training data
  4. Predict using model equation
- **Instance-Based**:
  1. Store all training data
  2. For *each* new instance:
     - Search entire dataset
     - Compute distances
     - Aggregate neighbors

Storage & Performance

Metric	Instance-Based	Model-Based
Storage Required	$O(n)$	$O(1)$
Prediction Speed	Slow	Fast
Scalability	Poor	Excellent
Handles Noise	Better	Requires regularization

5 When to Use

- **Instance-Based**:
  - Complex decision boundaries
  - Small datasets
  - Example: k-NN, kernel methods
- **Model-Based**:
  - Large datasets
  - Need for fast predictions
  - Interpretable patterns
  - Examples: Linear regression, neural networks