Artificial Intelligence and Machine Learning

Curated computer science study track (not a degree, no licensure claim). Original explanations and practice only, following the reviewed course handout.

The classical core, honestly measured

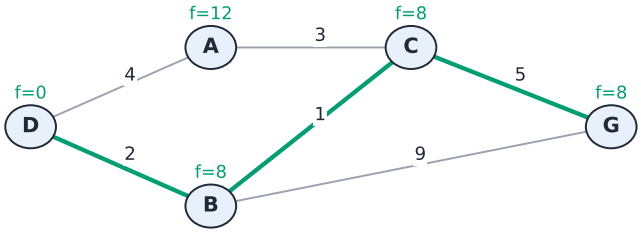
- 13.1-13.3 the spine: search, knowledge, probability; 13.4-13.5 learn from data.
- 13.6-13.7 keep it honest; the Overnight dataset (200 records, data card) feeds every number.

Search: f = g + h, and the estimate is the contract

1. **Tree** Nodes are states, edges are moves: BFS shortest when steps equal, DFS frugal; A* picks min g + h.
2. **Course** D-A 4, D-B 2, A-C 3, B-C 1, C-G 5, B-G 9 with h = A 8, B 6, C 5, G 0: pops D, B, C, G - optimal 8 via D-B-C-G, four expansions.
3. **Lie** Overestimate h(C) to 12: C carries f 15 vs Goal 11 - pops Goal at 11, suboptimal: the algorithm trusts a lie.

Search tree

A* expands four nodes, BFS visits all five



pop order:
D, B, C, G
optimal 8 via D-B-C-G

overestimating h(C)=12 pops G at 11: suboptimal

f = g + h with admissible h: the first goal pop is optimal

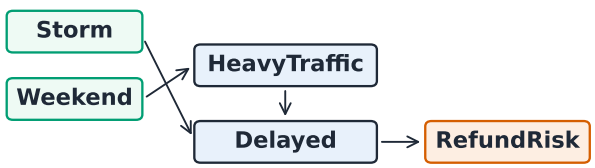
A* with an admissible h expands four nodes; BFS visits all five.

Knowledge and reasoning: rules make sound by construction

1. **KB** F1 Storm, F2 Weekend; R1 Storm-Delayed, R2 Weekend-HeavyTraffic, R3 HeavyTraffic-Delayed, R4 Delayed-RefundRisk.
2. **Waves** Forward: wave one Delayed and HeavyTraffic, wave two RefundRisk; backward for RefundRisk walks R4, R1 to a fact in three steps.
3. **Refute** Assume not-Q and resolve to the empty clause - the machinery behind SAT solvers and unit 10 chain.

Inference chain

two chains through the same rules



forward: facts (green) derive blue, then orange, in two waves



backward chaining touches only the relevant rules: R2, R3 never fire

modus ponens makes every derivation sound by construction

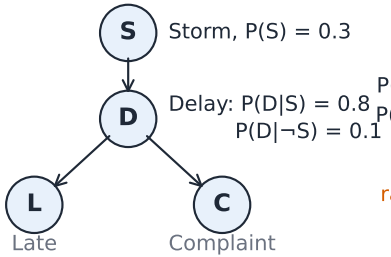
The backward chain from Delayed walks up to the known facts.

Probability models: evidence updates belief

1. **Numbers** P(S) 0.3, P(D|S) 0.8, P(D|notS) 0.1, P(C|D) 0.6, P(C|notD) 0.05: P(C|S) 0.49, P(C|notS) 0.105, P(C) 0.2205.
2. **Posterior** P(S|C) = 0.49 x 0.3 / 0.2205 = 0.667 - a complaint multiplies storm belief 2.5-fold; the no-storm leak is not negligible.
3. **Rare** With P(S) 0.03: posterior 0.126 - same evidence, belief collapsing to one-eighth; the normaliser and base rate carry the arithmetic.

Network graph

evidence flows up the arrows



the inference, in numbers:

$P(C|S) = 0.8 \times 0.6 + 0.2 \times 0.05 = 0.49$

$P(C|\neg S) = 0.1 \times 0.6 + 0.9 \times 0.05 = 0.105$

$P(C) = 0.49 \times 0.3 + 0.105 \times 0.7 = 0.2205$

$P(S|C) = 0.49 \times 0.3 \,/\, 0.2205 = 0.667$

rare storms (P = 0.03): posterior 0.126
the base rate is not a footnote

query P(Storm | Complaint): evidence explains backwards through the leak

Storm raises Delay; Delay explains Late and Complaint.

Regression and classification: the boundary and its bill

1. **Line** Points (1,2)(2,3)(3,5)(4,4)(5,6): w = 9/10, b = 1.3, residuals -0.2 -0.1 1.0 -0.9 0.2, SSE 1.9.
2. **Boundary** w1 2, w2 -1, b -1: the line x2 = 2x1 - 1; (2,2) scores +1 late, (2,5) scores -2 on time; margin = score magnitude.
3. **Overfit** 4th-degree poly: SSE 0 on train, about 42.7 held-out vs the line 0.55 - model choice by held-out error.

Decision boundary and data



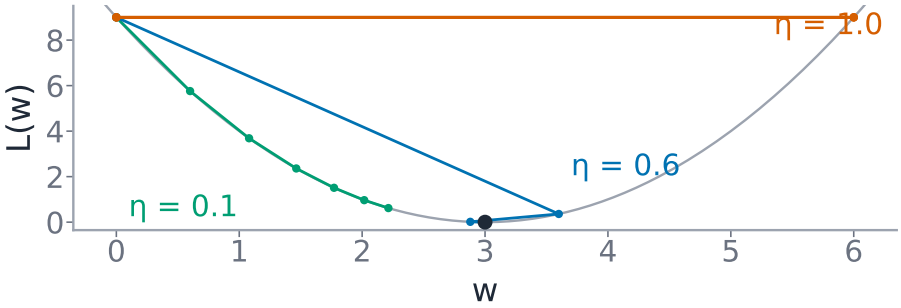
residuals -0.2, -0.1, +1.0, -0.9, +0.2 — SSE 1.9; a 4th-degree poly fits 0 here but held-out 42.7 vs the line 0.55

The fitted line with residuals; the boundary where the answer flips.

Optimisation and learning: descend, sized and scaled

1. **Quadratic** L(w) = (w-3)^2: eta 0.1 walks 0.6, 1.08, 1.464, 1.771, 2.017; eta 1 oscillates 0, 6 forever; stable iff eta < 1.
2. **Line** On the least-squares line: two steps from origin with eta 0.02 reach (0.84, 0.25); a thousand close on (0.9, 1.3).
3. **Ravine** L = (x1-1)^2 + 100(x2-1)^2: one rate cannot serve both curvatures - standardise features.

Loss surface and updates



minimum at $w = 3$; update $w \leftarrow w - 2\eta(w - 3)$; stable iff $\eta < 1$

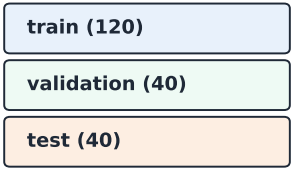
Three learning rates on the same bowl: crawl, damped, oscillation.

Evaluation and leakage: the honesty apparatus

1. **Matrix** TP 36, FP 10, FN 6, TN 148 on test 200: **accuracy** 0.92, **precision** 0.783, **recall** 0.857, F1 0.818 - base rate 21% makes 0.92 meaningful.
2. **Folds** 5-fold CV: 0.905 0.925 0.915 0.900 0.935 - mean 0.916 sd 0.013; a rival at 0.921 is within noise.
3. **Leak** A midnight status-code feature encodes the label: 0.99 train, 0.90 honest - the data card provenance column is the defence.

Split and confusion matrix

the split: test is spent once



5 folds over train+validation:
0.905 0.925 0.915 0.900 0.935
mean 0.916, sd 0.013
base rate 21% - the trivial predictor scores 0.79: 0.92 means something

the matrix, and its metrics

	pred late	pred on-time
true late	36	6
true on-time	10	148

accuracy **0.92**
precision **0.783**
recall **0.857**
F1 **0.818**

TP 36, FP 10, FN 6, TN 148

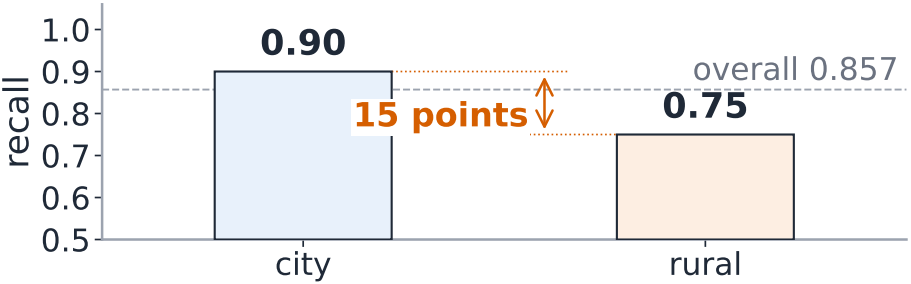
leakage is plumbing: a midnight status code
encodes the label - a score without folds is an anecdote

The split with test spent once; the matrix with its metrics.

Fairness and limitations: audit per group

1. **Gap** City **recall** $27/30 = 0.90$; rural $9/12 = 0.75$ - 15 points: the model misses 1 in 4 rural lates but 1 in 10 city lates.
2. **Cause** Rural lates are 13% of the region vs city 29% - the thinner class is under-served; fix records or weights, not the metric.
3. **Card** Overnight is synthetic, 200 records, two regions, one season: the card states intended use, the 15-point gap, out-of-scope uses.

Error disparity chart



city: TP 27, FN 3 (30 lates); rural: TP 9, FN 3 (12 lates)
lates are 13% of rural records, 29% of city records

Recall by region: city 0.90, rural 0.75.

Check yourself

- Course graph A* expansions before the goal…? (Four.)
- P(Storm | Complaint) with the course numbers…? (0.667.)
- The course line…? ($y = 0.9x + 1.3$.)
- City vs rural **recall**…? (0.90 vs 0.75.)