

Machine learning · supported focus

Supported focus · HL

Supported scope

- Supported HL focus. First assessment 2027 target; official PDF returns 403; older acquired brief is final assessment 2026. Remaining guide, assessment and practical requirements retain their recorded holds.
- State model conditions and distinguish fictional data from measurements.

Computing Data · 1.1

Reasoning

A model can score well by learning a shortcut in the training data. Evaluation must test the intended task on appropriate unseen examples. Start with a prediction. State the quantities or features you would compare, then decide what evidence could distinguish two explanations.

Computing Data · 2.1

Reasoning

A relational database uses tables, keys and relationships to organize data. Machine learning estimates patterns from data. Training and evaluation data serve different roles.

- primary key : An attribute set uniquely identifying a record;
- data leakage : Use of information unavailable at the intended prediction time.

Computing Data · 3.1

Reasoning

A primary key uniquely identifies a record; a foreign key relates records. In prediction, leakage can expose information unavailable at the real decision time. Accuracy alone may hide an unbalanced target distribution. Use fictional booking records to identify entities, attributes and relationships. For a learning exercise, use a public non-sensitive dataset, separate training and test data, and describe who may be affected by errors. The current 2027 CS objective scope awaits the acquired guide.

Computing Data · 4.1

Reasoning

State the known values and their units. Choose the relation because its assumptions fit this case, then rearrange before substitution. Known: a classifier makes 90 correct predictions among 100 cases. $\text{Accuracy} = \text{correct}/\text{total} \times 100 = 90\%$. If 90 cases belong to one class, always predicting that class gives the same accuracy and can still fail every minority-class case.

Computing Data · 4.2

Reasoning

A model makes 72 correct predictions out of 80. Find accuracy percentage. Use the same sequence: known quantities → model → relation → substitution → unit and interpretation.

Computing Data · 5.1

Reasoning

High accuracy is not evidence of fairness or causal understanding. A database primary key is not simply whichever field looks important. Return to the original observation. Explain what the result supports, which conditions it assumes, and one way to test a competing explanation.

Computing Data · 5.2

Reasoning

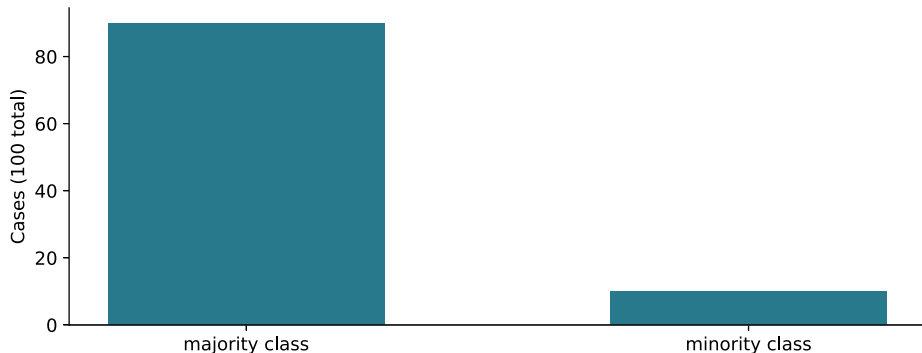
A classifier with 90% accuracy must treat every class equally well. This claim is false: High accuracy is not evidence of fairness or causal understanding. A database primary key is not simply whichever field looks important.

Computing Data · 5.3

Reasoning

Data models and responsible machine learning: A primary key uniquely identifies a record; a foreign key relates records. In prediction, leakage can expose information unavailable at the real decision time. Accuracy alone may hide an unbalanced target distribution.

Computing Data · diagram



- Original teaching diagram: Data models and responsible machine learning
- Use the labels and the stated model together.

Computing Data · original transfer

Your turn

A fictional dataset has 95 negatives and 5 positives. Compare an always-negative model with one finding 4 positives but making 10 false positives. Give accuracy and positive recall.

Computing Data · transfer reasoning

Method and interpretation

Baseline accuracy 95%, recall 0%. Second model has $TP=4, FN=1, FP=10, TN=85$: accuracy 89%, recall 80%. Accuracy alone hides minority performance; evaluation design and error costs matter.

Computing Data · executable trace

- Trace the state changes and boundary guard.
- Run the boundary case and compare the printed output.

```
tp, fn, fp, tn = 4, 1, 10, 85
total = tp + fn + fp + tn
accuracy = (tp + tn) / total
recall = tp / (tp + fn) if tp + fn else None
print(accuracy, recall)
positives = 0
print(0 / positives if positives else "UNDEFINED")
```

Check yourself

- Explain primary key using the stated model.
- Explain data leakage using the stated model.
- Change one value in the original transfer; explain which assumption still matters.