

4.1 Data models and responsible machine learning

Name: _____ Class: _____ Date: _____

Total: 10 marks

KEY WORDS primary key • data leakage

Objective

Data models and responsible machine learning. Supported HL concept preparation; original classroom practice, not an official IB paper or assessed learner work.

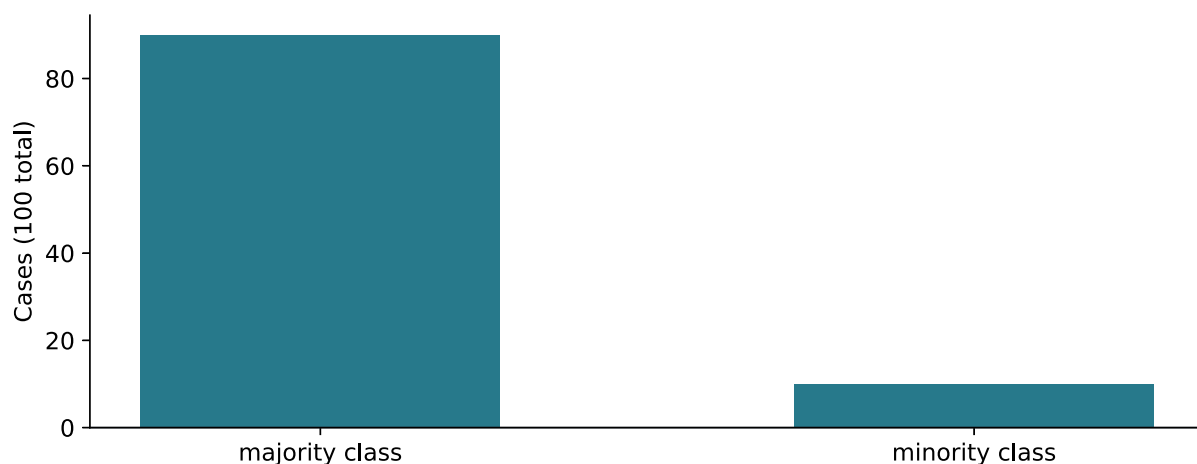
1 Key knowledge and worked examples

■ What would explain this observation?

- 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.

■ Build the model

- 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.



■ Choose evidence that can test it

- 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.

■ Work from known quantities

- 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.

Example:

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

■ Check the conclusion and its limits

- 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.

Warn:

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.

Key:

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.

■ Runnable trace and boundary

```
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")
```

Expected output:

```
0.89 0.8
UNDEFINED
```

Accuracy divides correct predictions by all cases; positive recall divides true positives by actual positives. With no positive cases, recall is undefined rather than zero. The dataset is fictional, with no trained model or personal data.

2 Practice

2.1 Why keep test data separate from training data? Choose one answer.

- **A** To assess generalization to unseen cases
- **B** To guarantee causation
- **C** To make all predictions correct

[1]

2.2 A model makes 72 correct predictions out of 80. Find accuracy percentage. Show the relation, substitution and units.

[2]

2.3 Decide whether this claim is true and justify: A classifier with 90% accuracy must treat every class equally well.

[2]

2.4 An attribute set uniquely identifying a record: write the technical term.

[1]

3 Exam-style questions

3.1 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. [4]
