AI Research Labs

Machine Learning Foundations · Practice test 4 · 55 questions

Logistic Regression and Classification

How to use this test. Answer all 55 questions without notes, then mark yourself with the answer key at the back. Each question has exactly one best answer. Score 75% or more and you've mastered the material; below that, reread the lessons for the questions you missed.

4.1 Regression versus Classification

Q1 concept Why is ordinary linear regression a poor choice for a yes/no outcome coded 0/1?

- A It always underfits whenever Y is binary
- B It cannot be fitted to 0/1 response values
- C It can only use one predictor at a time
- D Its predictions can fall below 0 or above 1

Q2 diagnose To predict which of three unordered product categories a customer buys, a team codes them 1, 2, 3 and fits linear regression. What's the core problem?

- A The predictions will be too accurate on the training data
- B Linear regression needs at least ten distinct target values
- C The categories should be coded 0, 1, 2 instead of 1, 2, 3
- D The coding imposes a meaningless order and equal spacing

Q3 concept Logistic regression is a linear model of what?

- A The probability of $Y = 1$
- B The odds of $Y = 1$
- C The class label Y itself
- D The log-odds of $Y = 1$

Q4 apply A logistic regression outputs $\hat{p} = 0.62$ for a customer. How does that become a class prediction?

- A Round each coefficient to 0 or 1, then recompute $\hat{p}$
- B The model outputs a class directly; 0.62 is its confidence
- C Pick the class whose coefficient has the larger value
- D Compare it with a threshold, e.g. predict 1 if $\hat{p} > 0.5$

Q5 apply You built a churn classifier. Which set of measures should you report?

- A The VIF of each predictor and the intercept
- B Sensitivity, specificity and AUC
- C R^2 and the RMSE of the 0/1 predictions
- D Adjusted R^2 and the coefficient p-values

4.2 Odds, Log-odds and the Sigmoid Function

Q6 `calculate` Convert the probability $p = 0.2$ to odds and to log-odds.

- A Odds 0.2, log-odds −1.61
- B Odds 0.8, log-odds −0.22
- C Odds 0.25, log-odds −1.39
- D Odds 4, log-odds 1.39

Q7 `calculate` The log-odds are 1.1. Using $e^{1.1} \approx 3$, what is the probability?

- A 0.75
- B 0.33
- C 0.25
- D 0.90

Q8 `calculate` A score is $z = -0.6$ (use $e^{0.6} \approx 1.82$). What is $p = \sigma(z)$?

- A 0.354
- B 0.549
- C 0.646
- D 0.400

Q9 `interpret` Odds of 3 for an event correspond to what probability?

- A 0.33
- B 0.67
- C 0.30
- D 0.75

Q10 `calculate` $p = 0.9$ has log-odds of about 2.2. What are the log-odds of $p = 0.1$?

- A About 0.1
- B About 0.22
- C About −0.9
- D About −2.2

4.3 The Logistic Regression Model

Q11 `calculate` A model has log-odds $= -3 + 0.5x_1 + 1.2x_2$. What is $P(Y = 1)$ at $(x_1, x_2) = (4, 1)$ (use $e^{-0.2} \approx 0.82$)?

- A 0.55
- B 0.20
- C 0.82
- D 0.45

Q12 `interpret` In that model, $e^{1.2} = 3.32$. How do you interpret it?

- A Each one-unit rise in x_2 multiplies the odds of $Y = 1$ by 3.32
- B Each one-unit rise in x_2 makes $Y = 1$ exactly 3.32 times as likely
- C Each one-unit rise in x_2 raises the probability of $Y = 1$ by 3.32
- D Each one-unit rise in x_2 adds 3.32 to the log-odds of $Y = 1$

Q13 concept If $\beta_1 = 0.7$, what does a one-unit increase in x_1 do?

A Adds 0.7 to the log-odds

B Multiplies the log-odds by 0.7

C Adds 0.7 to the probability

D Multiplies the probability by 0.7

Q14 interpret Two customers have churn probabilities 0.10 and 0.50. Each makes one more complaint, which has an odds ratio of 2.23. What happens?

A Both probabilities rise by the same amount, since the odds ratio is fixed

B The 0.10 customer's probability rises more, since it has more room to grow

C Both probabilities are multiplied by 2.23, giving 0.22 and 1.12

D Both odds rise ×2.23, but the 0.50 customer's probability rises more

Q15 concept Where is logistic regression's decision boundary at threshold 0.5?

A Where the predicted probability reaches exactly 1 everywhere

B Along the S-shaped sigmoid curve of predicted probabilities

C Where $\beta_0 + \beta^\top x = 0$, a straight line or hyperplane

D Along a circle centred on the mean of the training data

4.4 Maximum Likelihood Estimation

Q16 concept Why can't logistic regression be fitted with a closed-form formula like OLS's $(X^\top X)^{-1} X^\top y$?

A Its loss is not convex, so there are many local optima

B β sits inside the sigmoid, so the equations are nonlinear

C It has too many parameters to invert $X^\top X$ in practice

D Logistic regression doesn't have a loss function to minimise

Q17 calculate Labels $y = (1, 1, 0)$; predicted probabilities $p = (0.7, 0.4, 0.2)$. What is the log-likelihood?

A About −0.83

B About −2.10

C About −1.50

D About −0.51

Q18 concept Why do we maximise the log-likelihood rather than the likelihood itself?

A Taking logs prevents the model from overfitting the data

B The plain likelihood cannot be computed for logistic models

C The log gives a different, more accurate set of coefficients

D Logs turn products into sums, and the maximiser is unchanged

Q19 calculate A coefficient is $\hat{\beta} = -0.9$ with SE $= 0.45$. Is it significant at 5% (Wald test)?

A Yes: $z = -2$ and $|z| > 1.96$

B Yes: $|\hat{\beta}|$ is bigger than its standard error

C No: $z = -0.5$, which is too small

D No: a negative coefficient can't be significant

Q20 calculate A logistic model with no predictors is fitted to labels $(1, 0, 0, 1, 1)$. What is the maximum-likelihood estimate of p?

A 0.4
B 0.6
C 1.0
D 0.5

4.5 Cross-entropy / Log-loss

Q21 calculate For a case with true label 1, what is the cross-entropy loss when the model predicts 0.8, and when it predicts 0.2?

A 1.609 and 0.223
B 0.223 and 1.609
C 0.8 and 0.2
D 0.2 and 0.8

Q22 calculate For a case with true label 0, the model predicts $p = 0.1$ for class 1. What's the loss?

A 0.9
B 0.1
C 0.105
D 2.303

Q23 calculate Two cases: $y = (1, 0)$ with predictions $p = (0.9, 0.2)$. What is the mean cross-entropy?

A 1.609
B 0.164
C 0.150
D 0.328

Q24 concept Which loss does standard logistic regression minimise?

A Cross-entropy (negative log-likelihood)
B 0–1 loss, i.e. the misclassification rate
C Hinge loss, as in support vector machines
D Squared error on the 0/1 labels

Q25 interpret Two classifiers have identical accuracy, but A has a much lower cross-entropy than B. What does that tell you?

A A's probabilities are better: confident when right, cautious when wrong
B Nothing, because only accuracy matters once a threshold is fixed
C A is overfitting, because a lower loss means it memorised the data
D A must have used a better threshold than B to classify the cases

4.6 The Confusion Matrix

Q26 concept A fraud model flags a legitimate transaction as fraud. Which cell of the confusion matrix is that?

- A True negative
- B False negative
- C True positive
- D False positive

Q27 concept A screening test reports "no disease" for a patient who has the disease. Which cell is that?

- A False positive
- B False negative
- C True positive
- D True negative

Q28 calculate At threshold 0.5, four cases have (score, actual): (0.8, +), (0.6, −), (0.3, +), (0.2, −). What are TP, FP, FN and TN?

- A 2, 1, 0, 1
- B 1, 0, 1, 2
- C 2, 0, 0, 2
- D 1, 1, 1, 1

Q29 apply For an email spam filter, which error is usually worse?

- A A false negative: one spam email reaches the inbox
- B They always cost the same, so neither matters more
- C A false positive: a real email is hidden in spam
- D A true negative: the filter didn't learn anything

Q30 concept In hypothesis-testing language, a false positive is a...

- A Type III error
- B Type II error
- C Sampling error
- D Type I error

4.7 Accuracy, Precision, Recall and Specificity

Q31 calculate TP = 30, FN = 20, FP = 10, TN = 140. What are the sensitivity and specificity?

- A 0.60 and 0.933
- B 0.75 and 0.875
- C 0.85 and 0.933
- D 0.60 and 0.875

Q32 calculate Same table (TP = 30, FN = 20, FP = 10, TN = 140). What is the precision?

A 0.933
B 0.75
C 0.85
D 0.60

Q33 concept Which formula gives specificity?

A $TP/(TP + FP)$
B $TN/(TN + FP)$
C $TP/(TP + FN)$
D $FP/(FP + TN)$

Q34 interpret A test has sensitivity 0.95 and specificity 0.60. In plain words?

A It misses many sick people but rarely raises a false alarm
B It gives the correct answer for about 95% of all people tested
C It catches almost every sick person but flags 40% of healthy ones
D 60% of the people it flags as positive are actually sick

Q35 calculate A disease affects 1% of people. A test has 90% sensitivity and 90% specificity. Of the people who test positive, roughly what share actually have the disease?

A About 90%
B About 50%
C About 1%
D About 8%

4.8 Sensitivity and Specificity Trade-offs

Q36 interpret You lower the classification threshold from 0.5 to 0.2. What generally happens?

A Specificity rises and sensitivity falls
B Both sensitivity and specificity rise
C Sensitivity rises and specificity falls
D Neither changes; only accuracy does

Q37 concept With the same fitted model, can you raise sensitivity and specificity together by moving the threshold?

A Only if the model's accuracy is above 90%
B Yes, but only when the classes are balanced
C No: moving it trades one for the other
D Yes: there is always a threshold that improves both

Q38 interpret What happens at a threshold of 0, where every case is flagged positive?

A Sensitivity 0, specificity 1
B Sensitivity 1, specificity 0
C Accuracy becomes 1
D Both equal 0.5

Q39 diagnose A team lowers the threshold until sensitivity reaches 100%, then reports this as a big success. What's the catch?

- A Sensitivity can never genuinely reach 100% on real-world data
- B It means the model has overfitted the training data
- C Lowering the threshold also lowers the model's AUC
- D Flagging nearly everyone floods them with false positives

Q40 apply What would improve both sensitivity and specificity at the same time?

- A Lowering the threshold well below the 0.5 default
- B Raising the threshold well above the 0.5 default
- C Better features, so the two classes' scores overlap less
- D Oversampling the positive class in the test set before scoring

4.9 Threshold Selection and Classification Costs

Q41 calculate A false positive costs 1 and a false negative costs 9. At threshold 0.5 there are 5 FP and 10 FN; at 0.2 there are 25 FP and 3 FN. Which threshold has the lower cost?

- A 0.2, with cost 52 vs 95
- B They cost the same
- C 0.5, with cost 15 vs 28
- D 0.5, with cost 95 vs 52

Q42 calculate A false positive costs £500 and a false negative £1,500. With well-calibrated probabilities, which threshold minimises expected cost?

- A 0.50
- B 0.25
- C 0.33
- D 0.75

Q43 apply A missed diagnosis is 20 times as costly as a false alarm. Roughly where should the threshold be?

- A At 0.5, the default
- B Well above 0.5, around 0.95
- C Well below 0.5, around 0.05
- D Around 0.2, a typical value

Q44 diagnose A team picks the threshold that minimises cost on the test set and reports that cost as its expected cost in production. What's wrong?

- A They should always use 0.5 so the result is comparable to others
- B Costs can only be measured on the training set, never the test set
- C Nothing: costs, unlike accuracy, can't be overfitted at all
- D The threshold was tuned on the test set, so the cost is optimistic

Q45 concept If a false negative costs more than a false positive, which way should the threshold move from 0.5?

A It should stay at 0.5
B Down, so more cases are flagged
C Up, so fewer cases are flagged
D It depends only on accuracy

4.10 ROC Curve and AUC

Q46 concept What does an ROC curve plot?

A Precision vs recall, over all possible thresholds
B True positive rate vs false positive rate, over all thresholds
C Accuracy vs the threshold value, from 0 up to 1
D Specificity vs precision, measured at the default threshold of 0.5

Q47 calculate Positives are scored 0.8, 0.5, 0.4; negatives are scored 0.6, 0.3. What is the AUC, by counting pairs?

A 5/6 ≈ 0.83
B 3/6 = 0.50
C 4/6 ≈ 0.67
D 2/5 = 0.40

Q48 calculate There are 4 positives and 6 negatives. At some threshold, 3 positives and 2 negatives are flagged. Where is this point on the ROC curve?

A TPR 0.75, FPR 0.67
B TPR 0.33, FPR 0.75
C TPR 0.75, FPR 0.33
D TPR 0.60, FPR 0.40

Q49 interpret Model A has AUC 0.92 and model B has AUC 0.55. Which statement is best supported?

A B is useless at every threshold and must be scrapped
B A is correct for about 92% of the cases it scores
C A ranks positives above negatives much more reliably
D A's best threshold for classifying is exactly 0.5

Q50 concept Why is AUC called "threshold-independent", and what can't it tell you?

A It always uses 0.5, so it can't handle unequal costs
B It summarises all thresholds, so it can't say which one to use
C It ignores the scores, so it can't compare two models
D It uses only the best threshold, so it hides the trade-off

4.11 Class Imbalance

Q51 diagnose A fraud model reaches 99.5% accuracy on data where 0.5% of transactions are fraud. Why isn't this impressive?

- A Accuracy above 99% is a sure sign of overfitting
- B Predicting "not fraud" for everything also scores 99.5%
- C Accuracy can't be computed properly when classes are imbalanced
- D Fraud models need 100% accuracy to be deployable

Q52 diagnose "We oversampled the minority class, then split the data 80/20 for evaluation." What can go wrong?

- A The test set ends up too small to give a reliable score
- B The model becomes unable to predict the minority class on new data
- C Oversampling removes the majority class from the data
- D Copies of the same records land in both sets, inflating the score

Q53 concept In stratified oversampling, the training set is balanced. What should the validation set look like?

- A It should reuse the training cases
- B It should keep the original class ratio
- C It should be balanced 50/50 as well
- D It should contain only the rare class

Q54 interpret 95% of cases are negative. A model has 96% accuracy and 0.20 sensitivity. How good is it?

- A Excellent, because 96% accuracy is far above the 50% of chance
- B Barely better than the 95% baseline, and it misses 80% of positives
- C Good, because a sensitivity of 0.20 is quite normal for rare classes like this
- D Poor overall, because accuracy below 99% is not acceptable

Q55 concept What does giving the rare class a larger weight in the loss do?

- A Makes its errors costlier, so the model attends to it more
- B Moves the classification threshold to exactly 0.5
- C Removes most of the majority-class cases from training
- D Copies the rare cases into the validation set

Answer key

Each answer comes with the reasoning, including why the most tempting wrong option is wrong.

Q1 **D** · lesson 4.1

A straight line is unbounded, so its "probabilities" can be negative or exceed 1, and the constant-variance noise assumption fails for 0/1 data. It *can* be fitted with many predictors; the outputs just don't make sense.

Q2 **D** · lesson 4.1

Regression would treat category 3 as "further" from 1 than 2 is, which means nothing for unordered categories. Recoding to 0, 1, 2 has the same flaw. Use a classifier.

Q3 **D** · lesson 4.1

$\ln \frac{P}{1-P} = \beta_0 + \sum \beta_j X_j$. The probability itself is an S-shaped (sigmoid) function of the predictors, not a linear one.

Q4 **D** · lesson 4.1

Logistic regression gives a probability; a decision rule turns it into a class. The threshold is a choice (0.5 by default, lower when misses are costly).

Q5 **B** · lesson 4.1

Classifiers are judged by how they handle each class (sensitivity, specificity, precision) and how well they rank (AUC). R^2 and RMSE are regression measures, and VIFs describe predictors, not performance.

Q6 **C** · lesson 4.2

Odds $= 0.2/0.8 = 0.25$, and $\ln 0.25 = -1.386$. Odds 4 is the odds *against*; $-1.61 = \ln 0.2$ skips the odds step.

Q7 **A** · lesson 4.2

Log-odds of 1.1 means odds of $e^{1.1} \approx 3$, i.e. three to one in favour, so $p = 3/(1+3) = 0.75$. 0.33 wrongly treats the odds as $1/p$, and 0.25 is the probability of the other class.

Q8 **A** · lesson 4.2

$p = 1/(1+e^{0.6}) = 1/2.82 = 0.354$. A negative score always means a probability below 0.5. 0.646 is the answer for $z = +0.6$, and 0.549 is the odds $e^{-0.6}$ rather than the probability.

Q9 **D** · lesson 4.2

$p = \text{odds}/(1+\text{odds}) = 3/4$. "3 to 1 in favour" means 3 successes for every failure. 0.33 confuses odds with $1/p$.

Q10 **D** · lesson 4.2

Log-odds are symmetric around $p = 0.5$: $\ln(0.1/0.9) = -\ln 9 = -2.2$. Swapping p and $1-p$ just flips the sign.

Q11 **A** · lesson 4.3

$z = -3 + 2 + 1.2 = 0.2$, so $p = 1/(1+e^{-0.2}) = 1/1.82 = 0.55$; predicted class 1 at threshold 0.5. 0.20 is z itself, not a probability.

Q12 **A** · lesson 4.3

Exponentiated coefficients are odds ratios: they multiply the odds, holding x_1 fixed. The log-odds rise by 1.2 (not 3.32), and "3.32 times as likely" confuses odds with probability.

Q13 **A** · lesson 4.3

Coefficients act additively on the log-odds scale (and multiply the odds by $e^{0.7} \approx 2$). The change in probability depends on where you start.

Q14 **D** · lesson 4.3

Odds go 0.111 → 0.248 (p 0.10 → 0.20) and 1 → 2.23 (p 0.50 → 0.69). The same odds ratio gives +0.10 versus +0.19 in probability, because the sigmoid is steepest near 0.5.

Q15 **C** · lesson 4.3

$p > 0.5$ exactly when $z > 0$, so the boundary is $z = 0$: linear in x. That's why logistic regression is a linear classifier even though its probability curve is S-shaped.

Q16 **B** · lesson 4.4

Setting the gradient of the log-likelihood to zero gives equations with β inside $1/(1 + e^{-x^\top \beta})$, which can't be solved algebraically. The problem is convex (one optimum), so numerical methods find it reliably.

Q17 **C** · lesson 4.4

Probabilities of what happened: 0.7, 0.4 and $1 - 0.2 = 0.8$. $\ln 0.7 + \ln 0.4 + \ln 0.8 = -0.357 - 0.916 - 0.223 = -1.50$. −2.10 uses $\ln 0.2$ for the third case instead of $\ln 0.8$.

Q18 **D** · lesson 4.4

The log is strictly increasing, so the same β maximises both; sums of logs are just easier to differentiate and avoid tiny products. It changes neither the answer nor overfitting.

Q19 **A** · lesson 4.4

$z = \hat{\beta}/\mathrm{SE} = -0.9/0.45 = -2$. Significance depends on $|z|$, so negative coefficients count too. -0.5 divides the wrong way round, and being one SE from zero ($|z| > 1$) is not enough.

Q20 **B** · lesson 4.4

With a single shared probability, the MLE is the proportion of ones: 3/5. Setting the derivative of $3 \ln p + 2 \ln(1 - p)$ to zero gives the same answer. 0.4 is the proportion of zeros.

Q21 **B** · lesson 4.5

Loss is $-\ln p$ for the true class: $-\ln 0.8 = 0.223$, $-\ln 0.2 = 1.609$. Low probability on the truth is punished heavily; the loss isn't the probability itself.

Q22 **C** · lesson 4.5

For $y = 0$ the loss is $-\ln(1 - p) = -\ln 0.9 = 0.105$: the model gave 0.9 to the truth, so the loss is small. 2.303 would be right if the true label were 1.

Q23 **B** · lesson 4.5

Losses are $-\ln 0.9 = 0.105$ and $-\ln 0.8 = 0.223$. Their mean is $0.328/2 = 0.164$; 0.328 is the sum. 1.609 wrongly uses $-\ln 0.2$ for the second case.

Q24 **A** · lesson 4.5

Maximising the log-likelihood is the same as minimising cross-entropy. Hinge loss belongs to SVMs, and 0–1 loss is too jumpy to optimise directly.

Q25 **A** · lesson 4.5

Accuracy only checks which side of the threshold each prediction lands on. Cross-entropy scores the probabilities themselves, which matters for ranking, costs and calibration. It doesn't depend on any threshold.

Q26 **D** · lesson 4.6

Predicted positive (fraud), actually negative (legitimate): a false alarm. A false negative would be real fraud that got through.

Q27 **B** · lesson 4.6

Predicted negative, actually positive: a miss. In screening, misses are usually the costlier error.

Q28 **D** · lesson 4.6

Flagged (score > 0.5): 0.8 (+, TP) and 0.6 (−, FP). Not flagged: 0.3 (+, FN) and 0.2 (−, TN). So one of each.

Q29 **C** · lesson 4.6

Missing an important real email typically costs more than seeing one extra spam message, so spam filters favour fewer false positives. Which error is worse always depends on the use case.

Q30 **D** · lesson 4.6

A Type I error is rejecting a true null hypothesis (raising an alarm when nothing is there): a false positive. A Type II error is missing a real effect: a false negative.

Q31 **A** · lesson 4.7

Sensitivity $= 30/(30+20) = 0.60$; specificity $= 140/(140+10) = 0.933$. 0.75 is the precision $30/40$, and 0.85 the accuracy $170/200$.

Q32 **B** · lesson 4.7

Precision $= TP/(TP+FP) = 30/40 = 0.75$: of the 40 cases flagged, 30 were truly positive. 0.60 is the sensitivity (it divides by the actual positives instead).

Q33 **B** · lesson 4.7

Specificity is the share of actual negatives correctly cleared. $TP/(TP+FN)$ is sensitivity, $TP/(TP+FP)$ precision, and $FP/(FP+TN)$ the false positive rate ($1-$ specificity).

Q34 **C** · lesson 4.7

Sensitivity 0.95: few misses among the sick. Specificity 0.60: 40% of healthy people get a false positive. Neither number is overall accuracy or precision.

Q35 **D** · lesson 4.7

Per 1,000 people: 10 are sick and 9 test positive; 990 are healthy and 99 test positive. Precision $= 9/108 \approx 8\%$. With a rare condition, false positives swamp true ones even for a good test.

Q36 **C** · lesson 4.8

A lower threshold flags more cases: fewer positives are missed (sensitivity up) but more negatives are flagged (specificity down).

Q37 **C** · lesson 4.8

The threshold slides along a fixed trade-off (the ROC curve). Improving both needs a better model whose score distributions for the two classes overlap less.

Q38 **B** · lesson 4.8

Every positive is caught (no misses) and every negative is a false alarm. Accuracy then equals the share of positives, which is rarely useful.

Q39 **D** · lesson 4.8

Perfect sensitivity is easy if you flag everything. The cost appears as collapsed specificity and precision. Always report both sides of the trade-off. AUC doesn't depend on the threshold.

Q40 **C** · lesson 4.8

Moving the threshold trades one for the other. Only a better model, one whose scores separate the classes more cleanly, shifts the whole trade-off upward. Altering the test set changes measurements, not the model.

Q41 **A** · lesson 4.9

At 0.5: $5 \times 1 + 10 \times 9 = 95$. At 0.2: $25 \times 1 + 3 \times 9 = 52$. 15 and 28 just count errors without weighting them by cost.

Q42 **B** · lesson 4.9

Alert when $p > c_{FP}/(c_{FP}+c_{FN}) = 500/2000 = 0.25$. Misses cost more, so the threshold moves below 0.5. 0.75 swaps the two costs; 0.33 is $500/1500$.

Q43 **C** · lesson 4.9

$c_{FP}/(c_{FP}+c_{FN}) = 1/21 \approx 0.05$. Flag a case even when the probability is low, and monitor the resulting volume of false alarms.

Q44 **D** · lesson 4.9

The threshold is a hyperparameter. Choosing it on the test set makes the test cost a best-case number. Tune on validation data or by CV, then measure once on the test set.

Q45 **B** · lesson 4.9

Costly misses mean you'd rather raise more false alarms than miss positives, so you flag at lower probabilities. The theory gives $p > c_{FP}/(c_{FP}+c_{FN})$, which is below 0.5 here.

Q46 **B** · lesson 4.10

Each point is one threshold: sensitivity on the vertical axis, $1-$ specificity on the horizontal. The top-left corner is perfect, the diagonal is random guessing.

Q47 **C** · lesson 4.10

6 positive–negative pairs. 0.8 beats both negatives (2); 0.5 beats 0.3 only (1); 0.4 beats 0.3 only (1). That's 4 out of 6.

Q48 **C** · lesson 4.10

TPR $= 3/4 = 0.75$; FPR $= 2/6 = 0.33$. 0.60 and 0.40 divide by the 5 flagged cases instead of by the actual class sizes; 0.67 is the specificity.

Q49 **C** · lesson 4.10

AUC is the probability a random positive outscores a random negative: a ranking measure. It isn't accuracy, it doesn't pick a threshold, and 0.55 is weak but slightly better than chance.

Q50 **B** · lesson 4.10

AUC is the area under the whole ROC curve, so it doesn't depend on any one cut-off. To choose an operating threshold you still need costs or a target sensitivity.

Q51 **B** · lesson 4.11

The majority-class baseline already scores 99.5% while catching zero fraud. Report sensitivity, precision and a confusion matrix, and compare against that baseline.

Q52 **D** · lesson 4.11

Duplicated minority cases end up on both sides of the split, so the model is tested on records it has seen. Oversample only the training part, after splitting (inside each CV fold).

Q53 **B** · lesson 4.11

Validation should reflect the real world the model will face; otherwise precision and accuracy are meaningless. Only the training data is rebalanced, to help the model learn the rare class.

Q54 **B** · lesson 4.11

Predicting "negative" for everyone already gets 95%. The model adds one point of accuracy while catching only a fifth of the positives. Chance here is 95%, not 50%.

Q55 **A** · lesson 4.11

Class weights change how much each mistake counts during fitting, an alternative to oversampling that doesn't duplicate records. It doesn't delete data or touch the validation set.