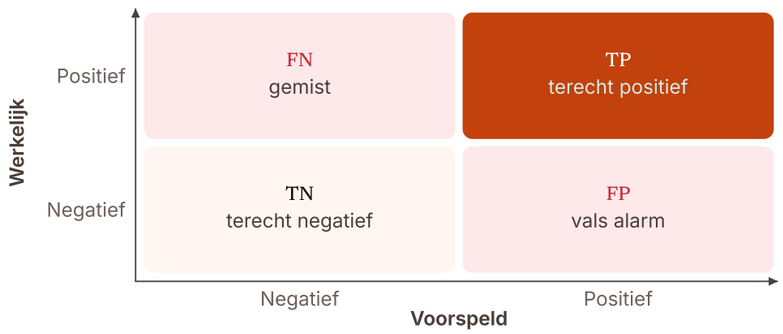
Logistische regressie, overfitting en validatie

POSTER · DE MODULE IN ÉÉN BLIK

Tentamen ma 7 dec 2026 · 180 min

1 Confusion matrix



Accuracy = (TP + TN) / n. Sensitivity = TP / (TP + FN). Specificity = TN / (TN + FP). Precision = TP / (TP + FP).

2 Overfitting voorkomen met validatie

1 Splits in train en test → 2 Fit op de train-set → 3 Voorspel de test-set → 4 Vergelijk train- en testprestatie

Beoordeel een model op data die het nog niet gezien heeft.

TENTAMENKANS

Logistische regressie (logitmodel)	98%
Confusion matrix, accuracy en sensitivity berekenen	97%
Accuracy is misleidend bij scheve klassen	95%
Logitcoëfficiënt is een verandering in log-odds, geen kans	94%
Odds en odds ratio (e^b)	93%
Overfitting en generalisatie	92%

VEELGEMAAKTE FOUTEN

- ✖ p = 0,06 is geen 6% kans op H0
- ✖ "2,75 keer zoveel kans" is fout: het zijn odds
- ✖ 0,5 is een gewoonte, geen regel
- ✖ p-value zegt niets over de grootte van het effect
- ✖ MLE-eigenschappen gelden voor grote n

ZO SCHRIJF JE HET ANTWOORD

1. Term model of begrip, Engelse vakterm

2. Definitie definitie of kenmerken, precies

3. Casusfeit concreet feit uit de casus

4. Conclusie oordeel, advies of gevolg

KERNBEGRIPPEN

logistic regression · binary dependent variable · log-odds · odds ratio · maximum likelihood · AIC · McFadden pseudo-R^2 · confusion matrix · accuracy · sensitivity · specificity · precision · threshold · ROC-curve · AUC · overfitting · cross-validation · bias-variance trade-off