PSY9511: Seminar 4

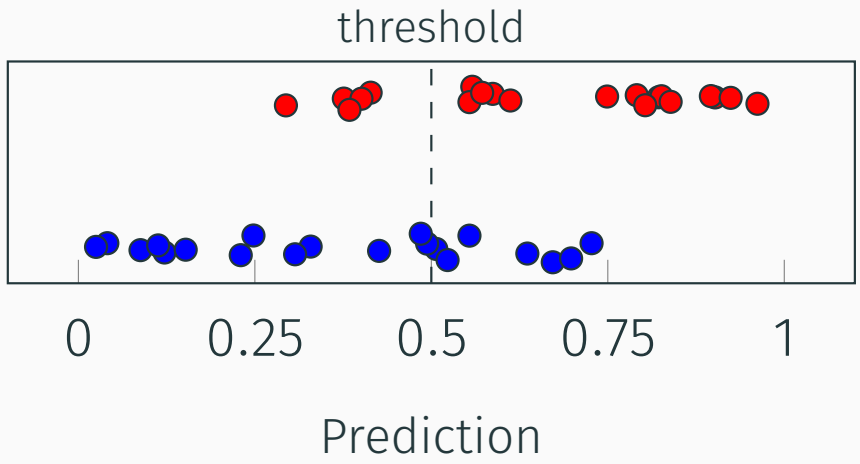
Model selection, validation and testing

Esten H. Leonardsen

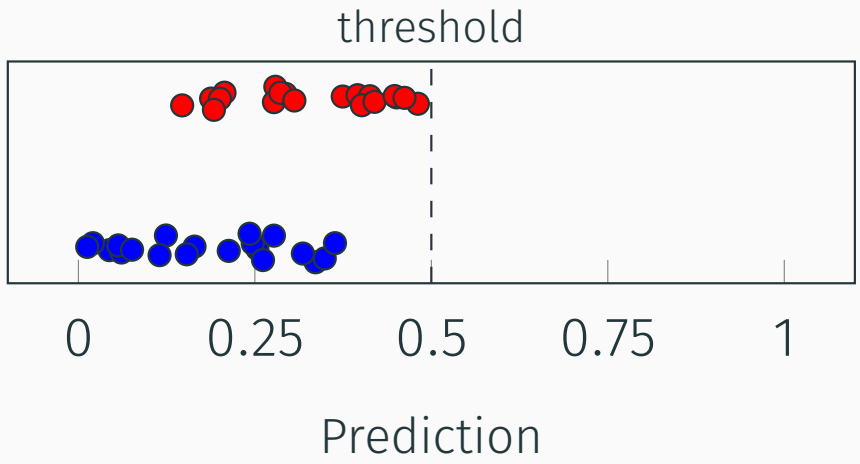
23.09.24

UNIVERSITETET
I OSLO

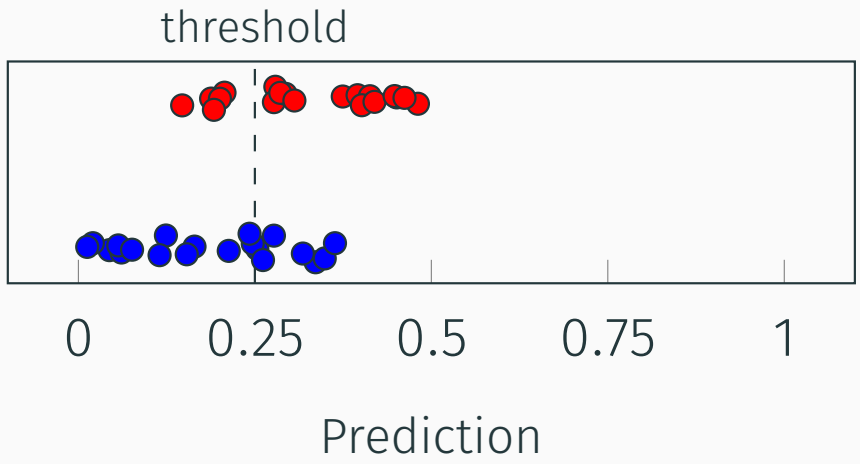
Classification metrics: Area under the curve



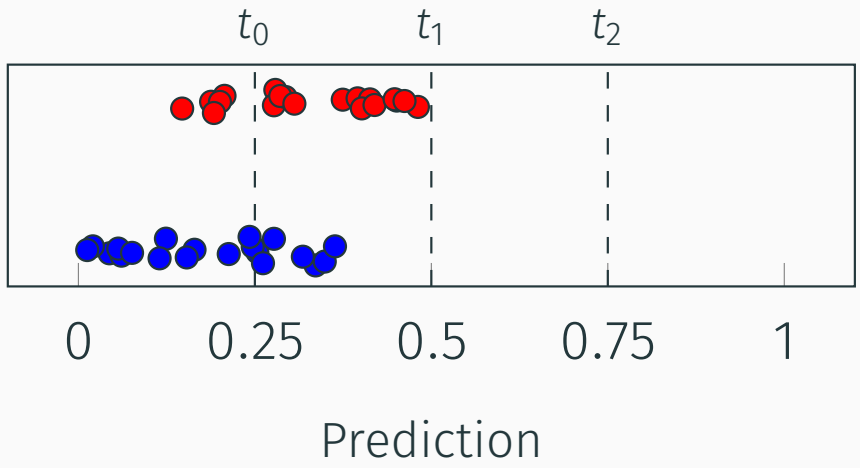
Classification metrics: Area under the curve



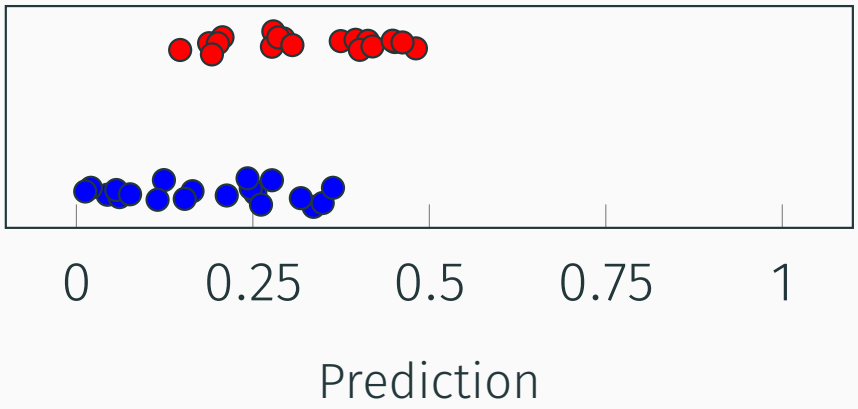
Classification metrics: Area under the curve



Classification metrics: Area under the curve

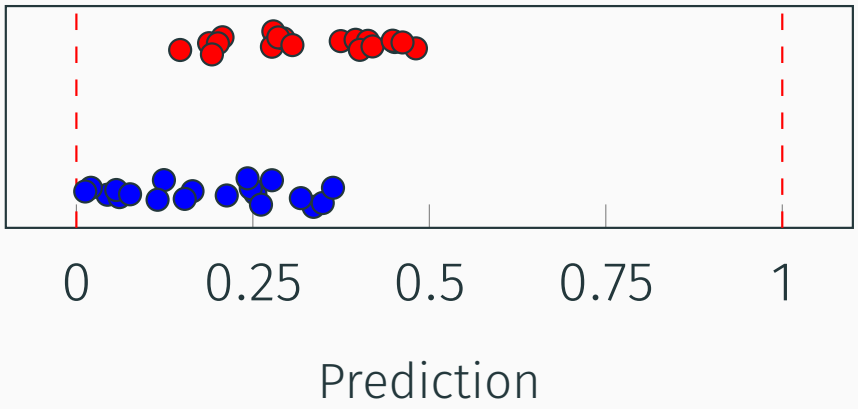


Classification metrics: Area under the curve



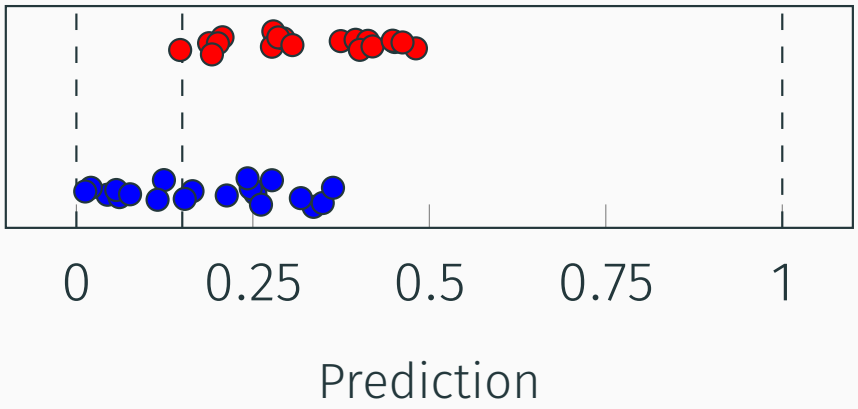
threshold	TPR	FPR

Classification metrics: Area under the curve



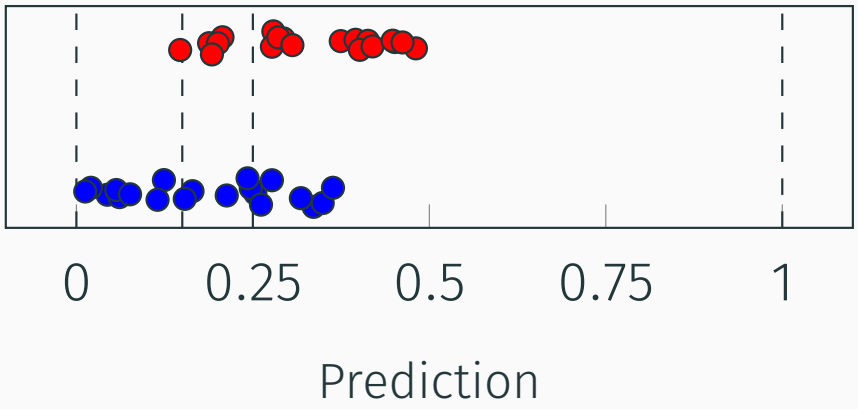
threshold	TPR	FPR
0	1	1
1	0	0

Classification metrics: Area under the curve



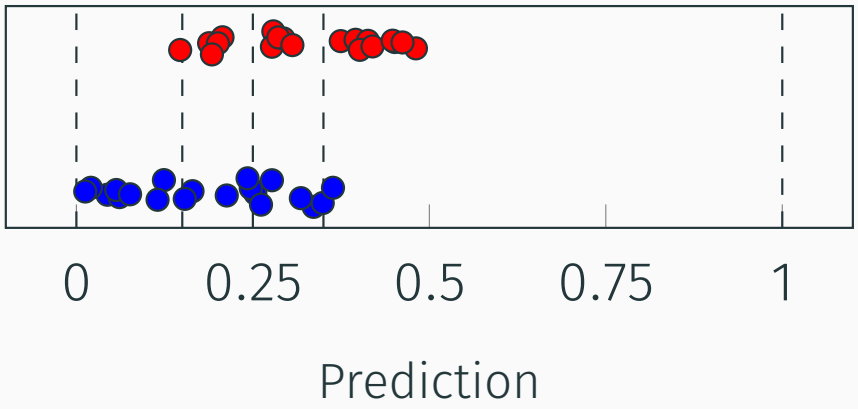
threshold	TPR	FPR
0	1	1
0.15	0.95	0.5
1	0	0

Classification metrics: Area under the curve



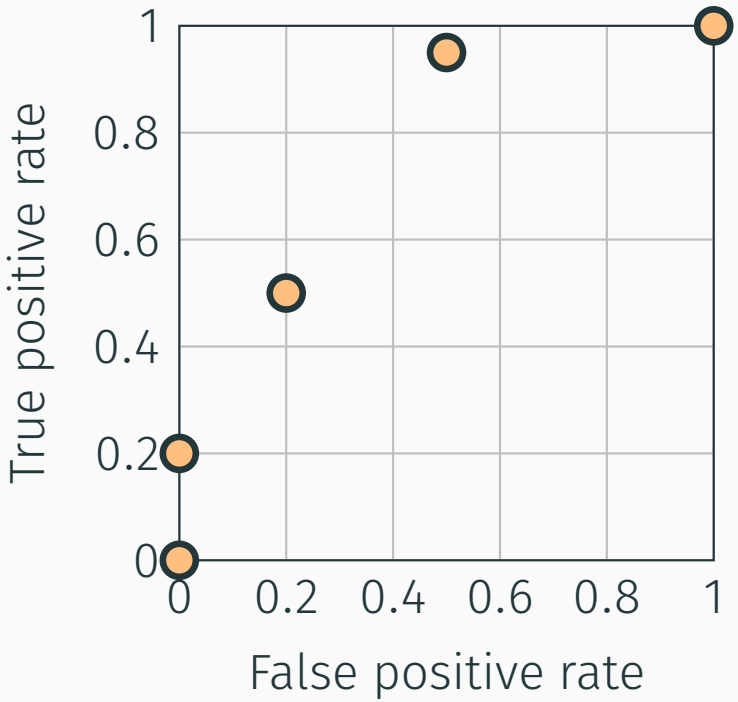
threshold	TPR	FPR
0	1	1
0.15	0.95	0.5
0.25	0.5	0.2
1	0	0

Classification metrics: Area under the curve



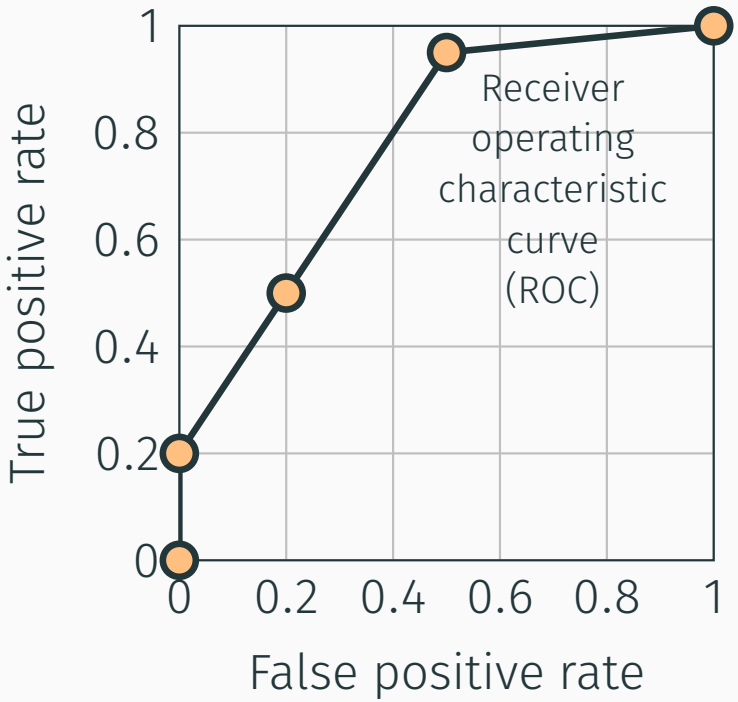
threshold	TPR	FPR
0	1	1
0.15	0.95	0.5
0.25	0.5	0.2
0.35	0.2	0.0
1	0	0

Classification metrics: Area under the curve



threshold	TPR	FPR
0	1	1
0.15	0.95	0.5
0.25	0.5	0.2
0.35	0.2	0.0
1	0	0

Classification metrics: Area under the curve

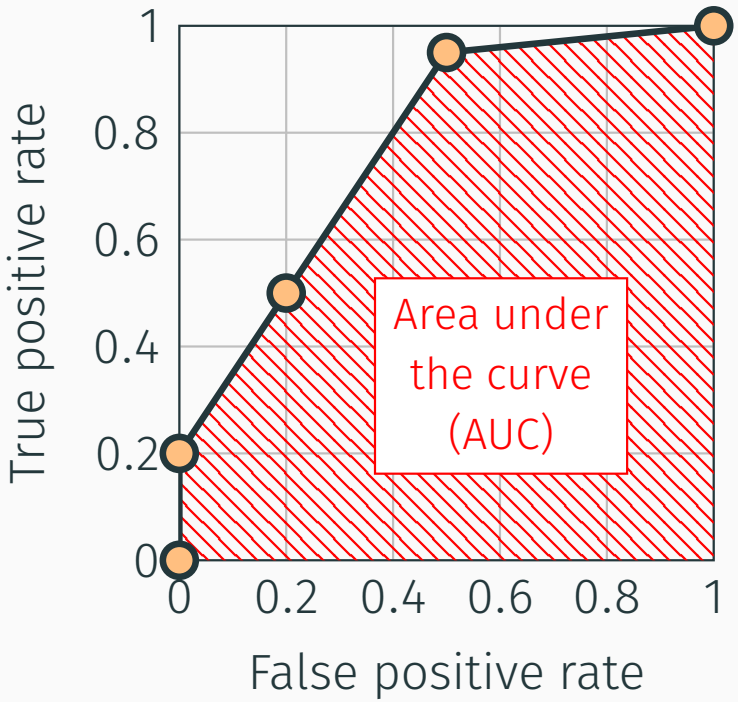


threshold	TPR	FPR
0	1	1
0.15	0.95	0.5
0.25	0.5	0.2
0.35	0.2	0.0
1	0	0

Classification metrics: Area under the curve

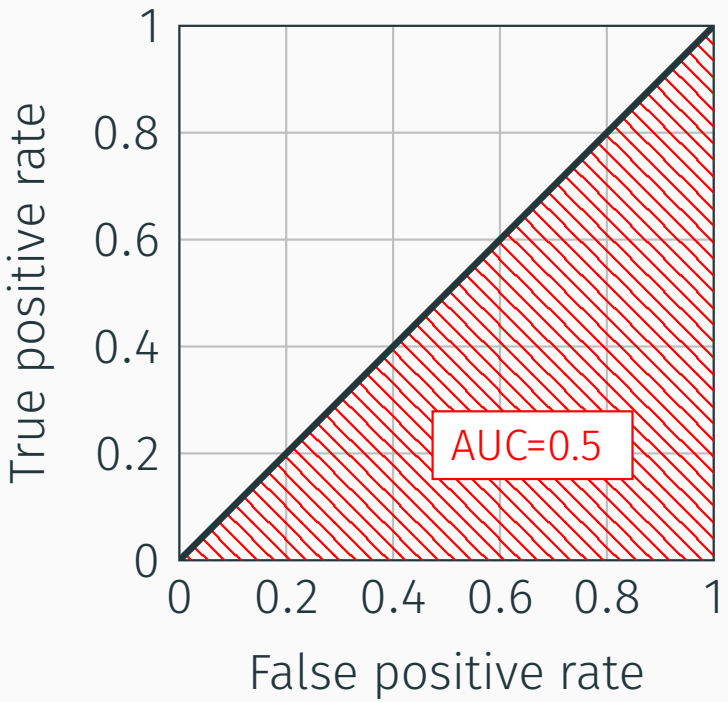
threshold	TPR	FPR
0	1	1
0.15	0.95	0.5
0.25	0.5	0.2
0.35	0.2	0.0
1	0	0

Classification metrics: Area under the curve



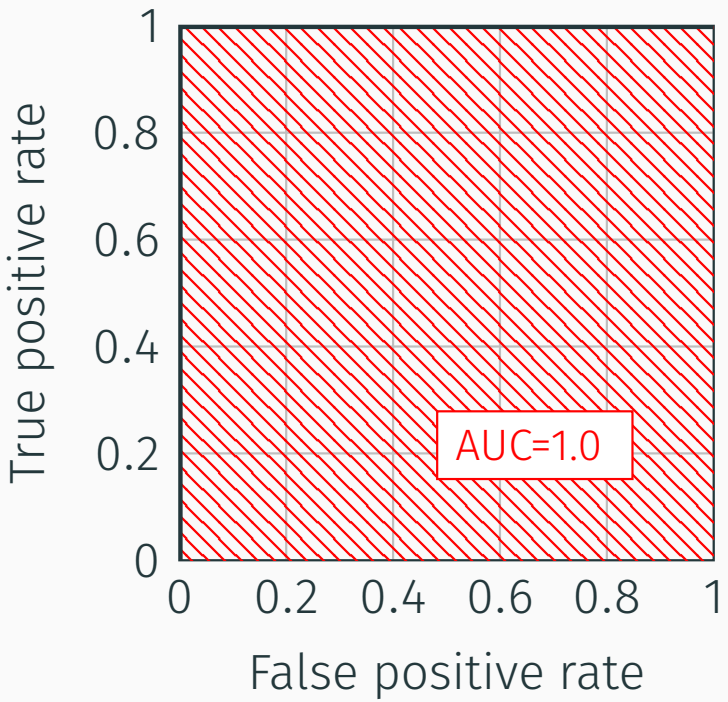
threshold	TPR	FPR
0	1	1
0.15	0.95	0.5
0.25	0.5	0.2
0.35	0.2	0.0
1	0	0

Classification metrics: Area under the curve



threshold	TPR	FPR
0	1	1
0.15	0.95	0.5
0.25	0.5	0.2
0.35	0.2	0.0
1	0	0

Classification metrics: Area under the curve



threshold	TPR	FPR
0	1	1
0.15	0.95	0.5
0.25	0.5	0.2
0.35	0.2	0.0
1	0	0

Classification metrics: Area under the curve

Area under the receiver operating characteristic curve (AUC/AUROC)

- A performance metric that does not rely on a correct classification threshold
- Measures whether the predictions are ranked correctly (e.g. patients have a higher prediction than controls)
- **Handles class imbalance** and is commonly used

Loss functions and performance metrics: Summary

Performance metrics and loss functions measure the performance of a predictive model

- There is a range of alternatives that can be used, each capturing a different aspect of a model's performance
- It is good practice to report more than one metric
- For regression:
  - MSE is a common loss function with nice mathematical properties.
  - MAE is an intuitive performance metric
- For classification:
  - Log-loss is the most common loss function for probabilistic classifiers
  - AUC is a widely used metric that is easy to interpret, handles class imbalance (to some degree), and is not reliant on the choice of classification threshold

Loss functions and performance metrics: Summary

http://localhost:8889/notebooks/notebooks%2FClassification%20metrics.ipynb