Curvas ROC y AUC

(Receiver Operating Characteristic and
Area Under the ROC Curve)

- Emanuel Rodríguez

A brief background

- Receiver-operating characteristic (ROC) analysis was originally developed during World War II to analyze classification accuracy in differentiating signal from noise in radar detection. Recently, the methodology has been adapted to several clinical areas heavily dependent on screening and diagnostic tests, in particular, laboratory testing, epidemiology, radiology, and bioinformatics [1].

Basic notions

- **ROC:** The ROC curve is a visual representation of model performance across different thresholds [5].
- **AUC:** The area under the ROC curve (AUC) represents the probability that the model, if given a randomly chosen positive and negative example, will rank the positive higher than the negative [5].

Motivation

- The motivation to use ROC is its ability to provide a comprehensive, threshold-independent evaluation of a model's performance, especially in scenarios where understanding trade-offs between sensitivity and specificity is critical. It is particularly valuable when comparing models or dealing with imbalanced datasets.

Motivation

- It is useful and insightful in interesting fields
- Provides a comprehensive view of how well the model performs beyond a single threshold, helping to identify its strengths and weaknesses.
- Different models might excel at different thresholds.
- A higher AUC generally indicates a better model overall.
- The ROC curve demonstrates the model's ability to distinguish between the positive and negative classes.
- The ROC curve focuses on **true positive rate** and **false positive rate**, providing insights into performance for both classes, regardless of imbalance.

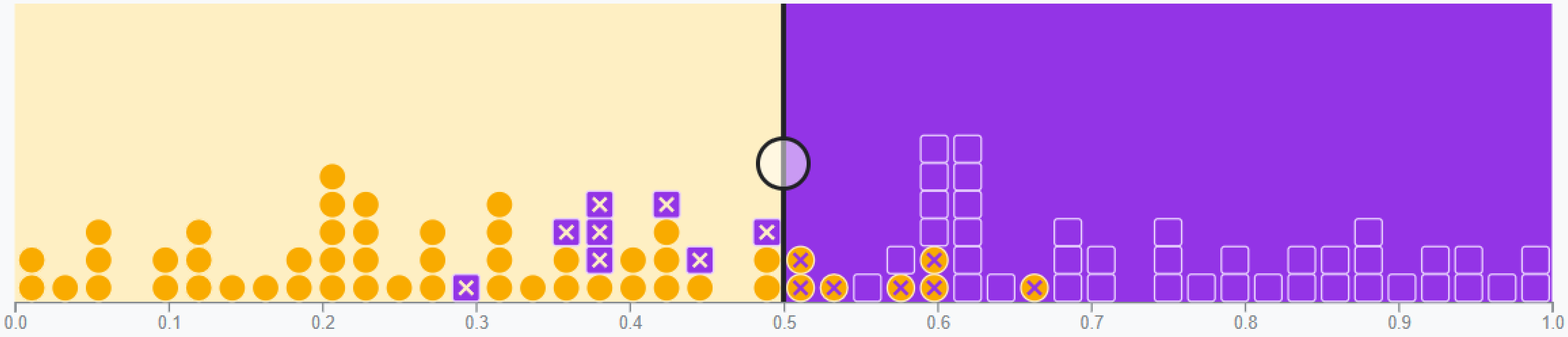
TPR AND FPR

	Predicted Positive	Predicted Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

- $TPR = Sensitivity = Recall = \frac{TP}{TP+FN}$
- $FPR = 1 - Specificity = 1 - \frac{TN}{TN+FP} = \frac{FP}{FP+TN}$

Threshold at 0.5

Classification threshold



Threshold: 0.50

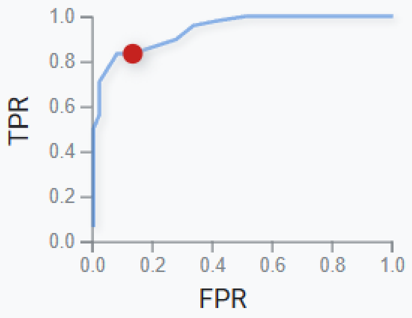
Confusion matrix

	Actually positive	Actually negative
Predicted positive	TP=40	FP=7
Predicted negative	FN=8	TN=44

Metrics

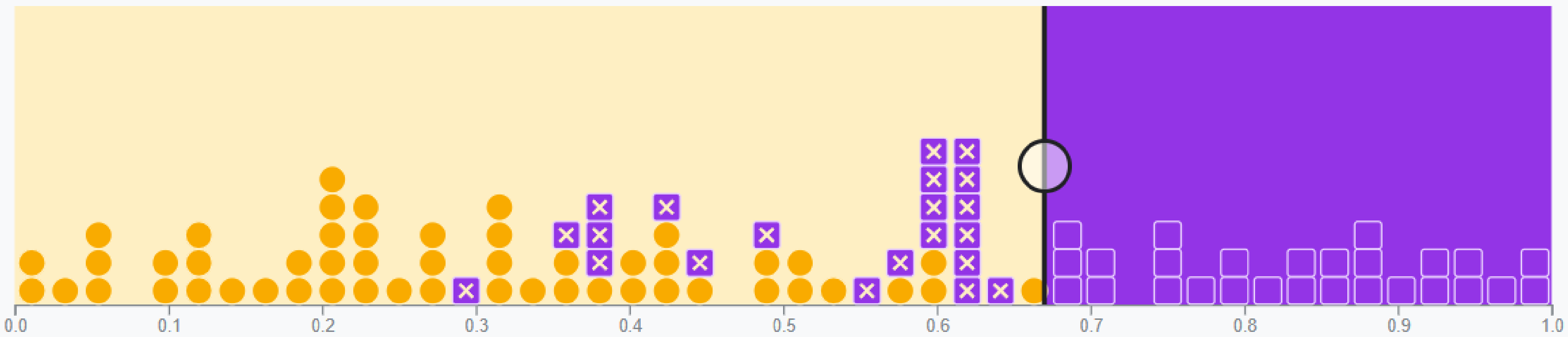
Accuracy	0.85
Precision	0.85
Recall	0.83

ROC curve



Avoiding False Positives

Classification threshold



Threshold: 0.67

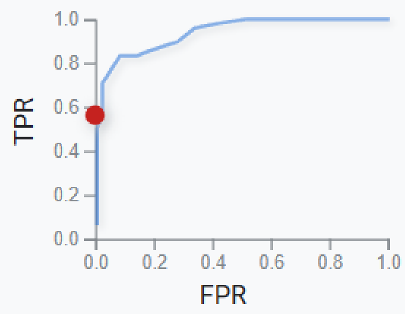
Confusion matrix

	Actually positive	Actually negative
Predicted positive	TP=27	FP=0
Predicted negative	FN=21	TN=51

Metrics

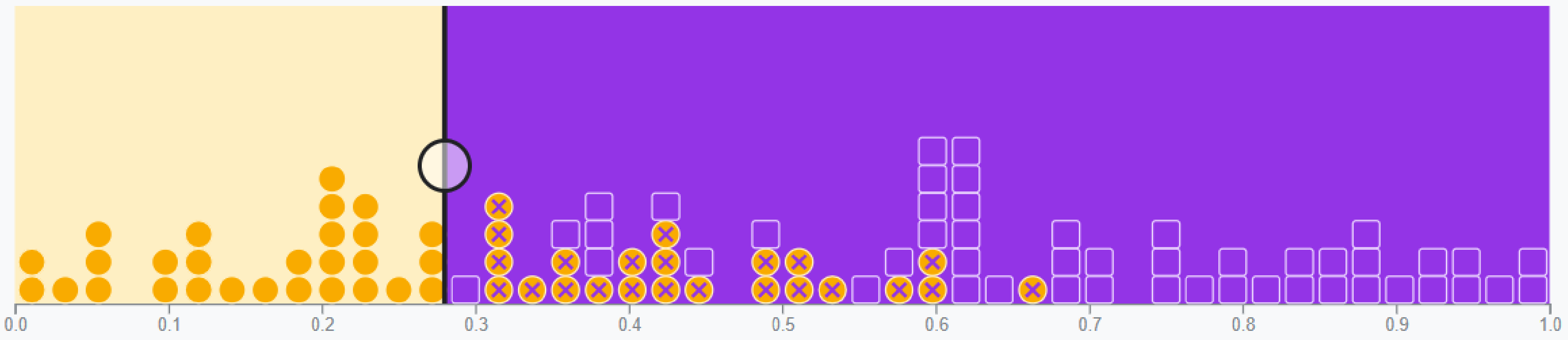
Accuracy	0.79
Precision	1.00
Recall	0.56

ROC curve



Avoiding False Negatives

Classification threshold



Threshold: 0.28

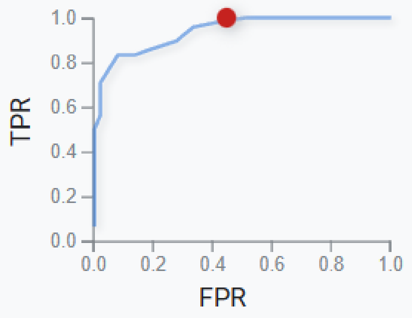
Confusion matrix

	Actually positive	Actually negative
Predicted positive	TP=48	FP=23
Predicted negative	FN=0	TN=28

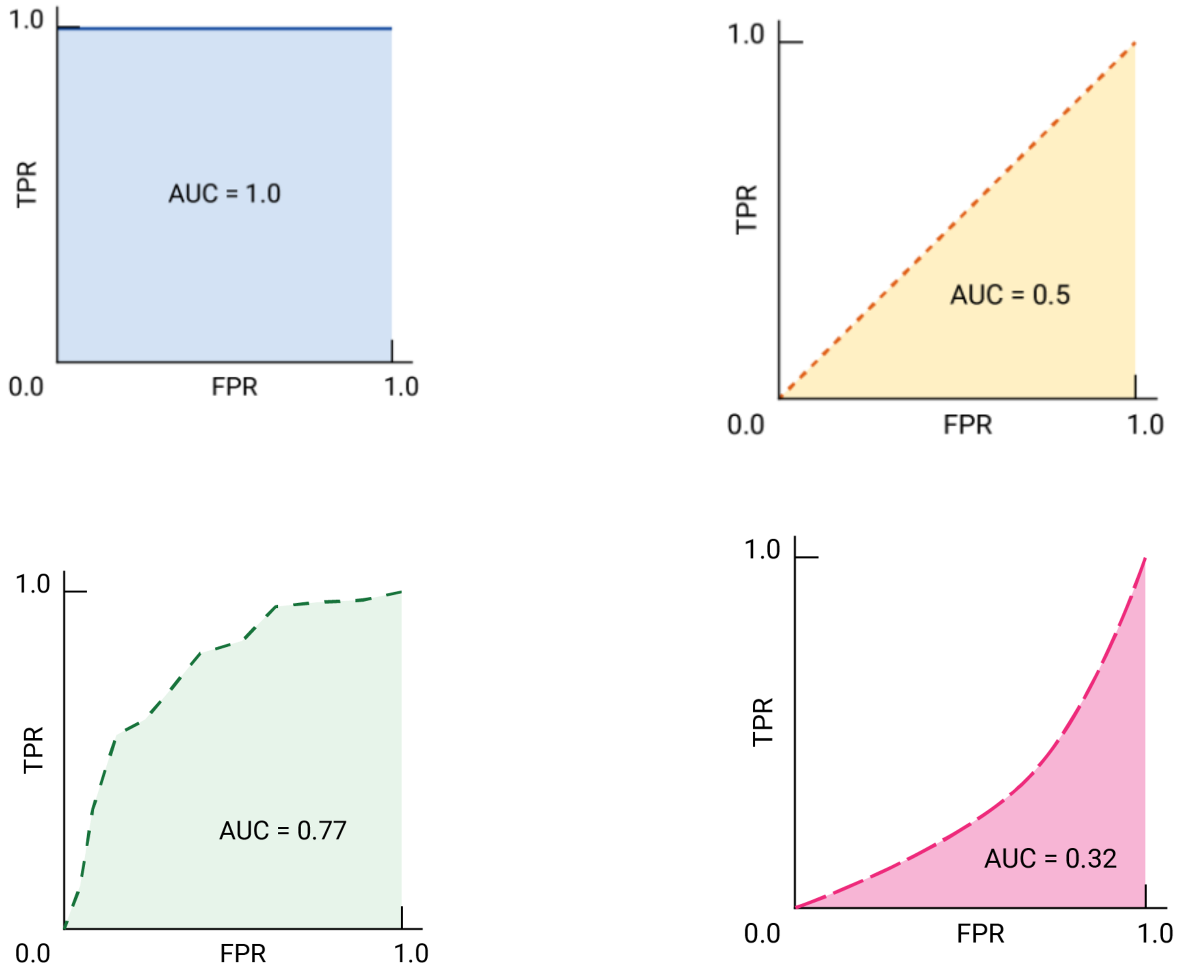
Metrics

Accuracy	0.77
Precision	0.68
Recall	1.00

ROC curve



Different ROC and AUC values



Regarding the Threshold

- A discrete classifier that returns only the predicted class gives a single point on the ROC space. But for probabilistic classifiers, which give a probability or score that reflects the degree to which an instance belongs to one class rather than another, we can create a curve by varying the threshold for the score [2].
- 2 common probabilistic classifiers:
  - Logistic Regression
  - MLP

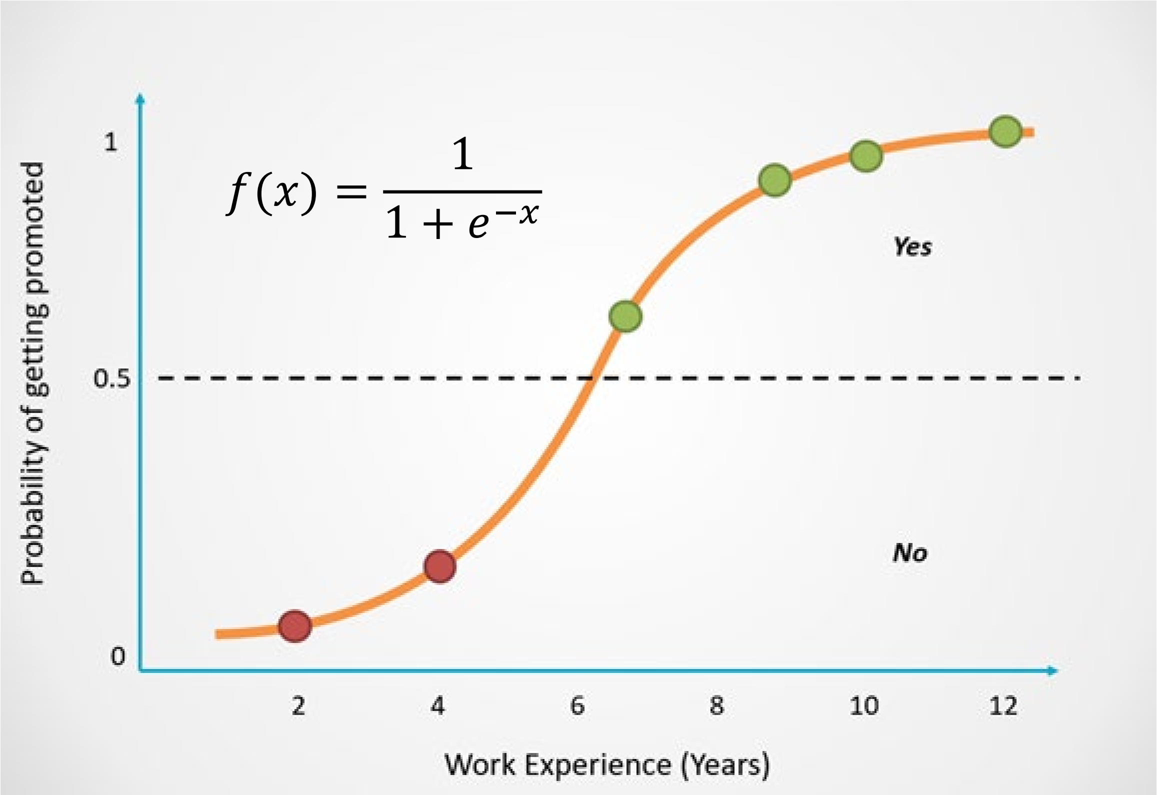
Experimental Setup

- Dataset: WDBC
- Validation method: LOO
- Algorithm: Logistic Regression
- Performance Measure: ROC & AUC

Dataset Info

- **Dataset: WDBC**
- Cardinality: 569 patterns
- Dimensionality: 30 attributes
- Type of attributes: Real, Integer
- Samples per class: 212 (M), 357(B)
- IR: 1.68 (Imbalanced)

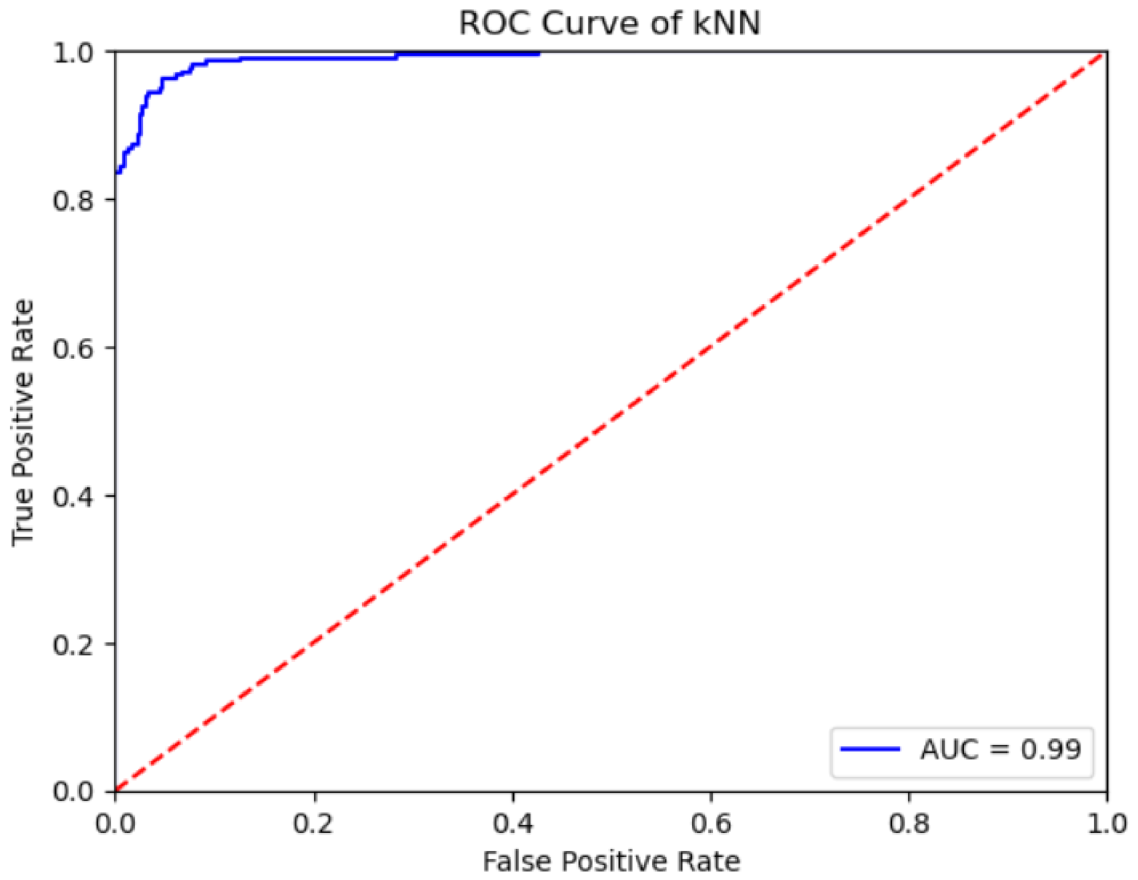
Logistic Regression



Results

FPR	TPR	TH
0.0	0.835	0.921
0.008	0.844	0.882
0.022	0.887	0.727
0.045	0.943	0.351
0.067	0.972	0.203
0.078	0.981	0.163
0.126	0.991	0.086
0.426	0.995	0.004
0.426	1.0	0.004

Results


ROC Curve of kNN
1.0
0.8
0.6
0.4
0.2
0.0
True Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate
AUC = 0.99

References

[1] Kelly H. Zou and A. James O'Malley and Laura Mauri, (2007), Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models, Circulation, 115, doi:10.1161/CIRCULATIONAHA.105.594929

[2] Carmen Chan, What is a ROC Curve and How to Interpret It, https://www.displayr.com/what-is-a-roc-curve-how-to-interpret-it/

[3] Aniruddha Bhandari,(2024), Guide to AUC ROC Curve in Machine Learning, https://www.analyticsvidhya.com/blog/2020/06/auc-roc-curve-machine-learning/

[4] Fan J, Upadhye S, Worster A., (2006), Understanding receiver operating characteristic (ROC) curves. Canadian Journal of Emergency Medicine, doi:10.1017/S1481803500013336

[5] Google, Machine Learning Crash Course. Classification: ROC and AUC, https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc