Linear Regression

Logistic Regression

**What Are ROC Curves?**

A useful tool when predicting the probability of a binary outcome is the [Receiver Operating Characteristic curve](https://en.wikipedia.org/wiki/Receiver_operating_characteristic), or ROC curve.

It is a plot of the false positive rate (x-axis) versus the true positive rate (y-axis) for a number of different candidate threshold values between 0.0 and 1.0. Put another way, it plots the false alarm rate versus the hit rate.

The **true positive rate** is calculated as the number of true positives **divided by the sum of the number of true positives and the number of false negatives**. It describes **how good the model is at predicting the positive class when the actual outcome is positive.**



|  |  |
| --- | --- |
| 1 | **True Positive Rate = True Positives / (True Positives + False Negatives)** |

The true positive rate is also referred to as sensitivity.



|  |  |
| --- | --- |
| 1 | **Sensitivity = True Positives / (True Positives + False Negatives)** |

**The false positive rate is calculated as the number of false positives divided by the sum of the number of false positives and the number of true negatives.**

It is also called the false alarm rate as it summarizes **how often a positive class is predicted when the actual outcome is negative.**



|  |  |
| --- | --- |
| 1 | False Positive Rate = False Positives / (False Positives + True Negatives) |

The false positive rate is also referred to as the inverted specificity where **specificity is the total number of true negatives divided by the sum of the number of true negatives** and false positives.



|  |  |
| --- | --- |
| 1 | Specificity = True Negatives / (True Negatives + False Positives) |

Where:



|  |  |
| --- | --- |
| 1 | False Positive Rate = 1 – Specificity |

The ROC curve is a useful tool for a few reasons:

* The curves of different models can be compared directly in general or for different thresholds.
* The area under the curve (AUC) can be used as a summary of the model skill.

The shape of the curve contains a lot of information, including what we might care about most for a problem, the expected false positive rate, and the false negative rate.

To make this clear:

* Smaller values on the x-axis of the plot indicate lower false positives and higher true negatives.
* Larger values on the y-axis of the plot indicate higher true positives and lower false negatives.

If you are confused, remember, when we predict a binary outcome, it is either a correct prediction (true positive) or not (false positive). There is a tension between these options, the same with true negative and false negative.

A skilful model will assign a higher probability to a randomly chosen real positive occurrence than a negative occurrence on average. This is what we mean when we say that the model has skill. Generally, skilful models are represented by curves that bow up to the top left of the plot.

A no-skill classifier is one that cannot discriminate between the classes and would predict a random class or a constant class in all cases. A model with no skill is represented at the point (0.5, 0.5). A model with no skill at each threshold is represented by a diagonal line from the bottom left of the plot to the top right and has an AUC of 0.5.

A model with perfect skill is represented at a point (0,1). A model with perfect skill is represented by a line that travels from the bottom left of the plot to the top left and then across the top to the top right.

An operator may plot the ROC curve for the final model and choose a threshold that gives a desirable balance between the false positives and false negatives.

Intuitively, we know that proclaiming all data points as negative in the terrorist detection problem is not helpful and, instead, we should focus on identifying the positive cases. The metric our intuition tells us we should maximize is known in statistics as [**recall**](https://en.wikipedia.org/wiki/Precision_and_recall), or the ability of a model to find all the relevant cases within a dataset.

 Recall can be thought as of a model’s ability to find all the data points of interest in a dataset.

https://miro.medium.com/max/60/1*gscG4JdjnyU5QkqNDqBg_w.png?q=20

this new model would suffer from low [**precision**](https://en.wikipedia.org/wiki/Precision_and_recall), or the ability of a classification model to identify only the relevant data points.

While recall expresses the ability to find all relevant instances in a dataset, precision expresses the proportion of the data points our model says was relevant actually were relevant.

