# Performance metrics – Intro

1. Hyperparam tuning: Search

* A search consists of:
* Hyperparam space
* A method for sampling candidate hyperparams
* A cross validation scheme
* **A performance metric to minimize (or maximize)**

1. Metrics

* Classification metrics
* Regression metrics
* Metrics in scikit-learn and how they are used
* Creating our own metrics

# Classification metrics

1. Overview
2. Accuracy

* Percentage or fraction of correct predictions
* Fraction of the predictions that the model got right
* Confusion matrix

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

* TN: # of negative samples correctly classified – True negatives
* FP: # of negative samples incorrectly classified as positive – False negatives
* TP: # of positive samples correctly classified – True positives
* FP: # of positive samples incorrectly classified as negative – False positives

1. Precision and Recall

* TP rate (recall or sensitivity) – How many observations from class 1 are going to be correctly identified by the model
* Positive predictive value (precision) – percentage of observations that are correctly classified from all those classified as members of class 1
* F1 score - weighted harmonic mean of precision and recall

1. FPR and FNR

* False positive rate
* False negative rate

1. Receiving Operating Characteristic Curve (ROC curve)

* Plots benefits (TPR) vs. costs (FPR) at different thresholds
* ROC-AUC: area under the ROC curve

1. Loss function

# Regression metrics

1. MSE, RMSE, and MAE
2. R squared

* How much of the total variance that exists in our data is explained by the model
* R2 takes values from 0 to 1

# Scikit-learn metrics

1. Metric optimization

* Classification
* Maximize:
  + ROC-AUC
  + Accuracy
  + Precision, Recall
  + F1-score
* Minimize
  + FPR, FNR, log-loss
* Regression
* Maximize:
  + R squared
* Minimize
  + MSE, RMSE, MAE

1. Scikit-learn optimization

* GridSearch and RandomSearch
* Cross\_validate
* To minimize the required metric, it **maximizes the negated version** of the metric
* Scikit-learn available metrics (see slides for link)
* Not all metrics available, but we can make our own metrics with the make\_scorer