UNIVERSITÄT
ZU KÖLN

MACHINE LEARNING EVALUATION

Sprachverarbeitung: Übung

Janis Pagel

Institut für Digital Humanities

13 May 2025

01

SOLUTION TO EXERCISE 04

Solution to Exercise 04

- https://lehre.idh.uni-koeln.de/site/assets/files/5615/solution04.pdf

02

RECAP

Recap

- Evaluation for Machine Learning Predictions
- Types of Errors
  - True Positives, True Negatives, False Positives, False Negatives
- Accuracy
- Precision and Recall
- F-Measure
- Macro and Micro Average

03

EVALUATION WITH SCIKIT-LEARN

scikit-learn

- Python library that implements many machine learning algorithms
- Can also be used for automatic calculation of different evaluation metrics
  - The sub-library `sklearn.metrics` implements the evaluation scores

sklearn.metrics

```
import sklearn.metrics
import pandas as pd
df = pd.DataFrame(
    {"Gold": [1, 1, 0, 1, 0, 0],
    "System": [1, 1, 0, 1, 1, 0]}
)

print(sklearn.metrics.accuracy_score(y_true = df["Gold"], y_pred = df["System"]))
~> 0.8333333333333334
print(sklearn.metrics.precision_score(y_true = df["Gold"], y_pred = df["System"]))
~> 0.75
print(sklearn.metrics.recall_score(y_true = df["Gold"], y_pred = df["System"]))
~> 1.0
print(sklearn.metrics.f1_score(y_true = df["Gold"], y_pred = df["System"]))
~> 0.8571428571428571
print(sklearn.metrics.classification_report(y_true = df["Gold"], y_pred = df["System"]))
~>               precision    recall  f1-score   support
~>
~>            0       1.00      0.67      0.80         3
~>            1       0.75      1.00      0.86         3
~>
~>     accuracy                           0.83         6
~>    macro avg       0.88      0.83      0.83         6
~> weighted avg       0.88      0.83      0.83         6
# scikit-learn calls micro average "weighted average"
```

04

EXERCISE 05

Exercise 05

- https://lehre.idh.uni-koeln.de/site/assets/files/5615/exercise05.pdf

UNIVERSITY
OF COLOGNE

Janis Pagel
Institut für Digital Humanities

eMail janis.pagel@uni-koeln.de
Homepage https://janispagel.de
Phone +49 221 470 5749

Foto: Gregor Hübl