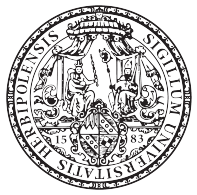

Introduction to Informatics
WiSe 2024/2025

Prof. Dr. Ingo Scholtes
Chair of Informatics XV
University of Würzburg

Exercise Sheet 06

Published: December 3, 2024
Due: December 16, 2024

Please upload your solutions to WueCampus as a scanned document (image format or pdf), a typesetted PDF document, and/or as a `jupyter` notebook.

1. **Confusion Matrix**

   You are given the ground truth labels and predicted labels for a binary classification problem.

   - Ground truth labels: `[1, 0, 1, 1, 0, 1, 0, 0, 1, 0]`
   - predicted labels: `[1, 0, 0, 1, 0, 1, 1, 0, 1, 0]`

   (a) Calculate the confusion matrix.

   (b) Calculate the accuracy, precision, recall, and F1 score.

2. **Logistic Regression**

   (a) The provided notebook implements a logistic regression model on the moons dataset from scikit-learn. Review the code and complete the missing cells to ensure that the model is properly trained, evaluated, and visualized.

   (b) Run the code again, but this time use the **circle dataset** instead of the moons dataset. Use the `make_circles` function from scikit-learn with the following parameters:

   - `n_samples=300`
   - `noise=0.05`
   - `factor=0.5`
   - `random_state=42`

   Adapt the code as needed to train, evaluate, and visualize the model on the circle dataset.

   (c) Compare the evaluation of the two models (moons dataset and circle dataset).

   - Which model has better performance, and why?
   - Discuss the limitations of the logistic regression model for these datasets.