

QuaPy: A Publicly Available Python-Based Software Library for Quantification

Alejandro Moreo, Andrea Esuli and Fabrizio Sebastiani

Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, 56124 Pisa, Italy

Abstract

QuaPy (<https://github.com/HLT-ISTI/QuaPy>) is an open-source framework for performing quantification, written in Python. QuaPy provides implementations of a number of baseline methods and advanced quantification methods (including CC, ACC, PCC, PACC, X, MAX, T50, MS, SLD, HDy, SVM(KLD), SVM(NKLD), SVM(Q), SVM(AE), SVM(RAE), QuaNet, and ensemble methods), of routines for quantification-oriented model selection, of several broadly accepted evaluation measures (including AE, RAE, SE, KLD, NKLD), and of robust evaluation protocols (including the generation of test samples by sampling uniformly at random from all legitimate prevalence distributions) routinely used in the field. QuaPy also makes available datasets commonly used for testing quantifiers, and offers visualization tools (including diagonal plots, error-by-shift plots, and bias boxplots) for facilitating the analysis and interpretation of the results. The software is open-source, publicly available under a BSD-3 licence via GitHub, can be installed via `pip` (<https://pypi.org/project/QuaPy/>), and is open to contributions of new material by the quantification community.

LQ 2021: 1st International Workshop on Learning to Quantify, Gold Coast, AU, November 1 and November 5, 2021.

✉ alejandro.moreo@isti.cnr.it (A. Moreo); andrea.esuli@isti.cnr.it (A. Esuli); fabrizio.sebastiani@isti.cnr.it (F. Sebastiani)

🆔 0000-0002-0377-1025 (A. Moreo); 0000-0002-5725-4322 (A. Esuli); 0000-0003-4221-6427 (F. Sebastiani)



© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)