Tutorial Proposal: "Conformal prediction: basics and selected recent topics"

1 Organisers

- **Name:** Sébastien Destercke
  Affiliation: Heudiasyc UMR CNRS 7253
  Email: sebastien.destercke@hds.utc.fr
  Webpage: `https://www.hds.utc.fr/~sdesterc/dokuwiki/`

- **Name:** Soundouss Messoudi
  Affiliation: UTC, Heudiasyc UMR CNRS 7253
  Email: soundouss.messoudi@hds.utc.fr
  Webpage: `https://www.hds.utc.fr/~smessoud/dokuwiki/`

- **Name:** Bruce Cyusa Mukama
  Affiliation: UTC, Heudiasyc UMR CNRS 7253
  Email: bruce.cyusa-mukama@hds.utc.fr
  Webpage: `https://www.hds.utc.fr/~bcyusamu/dokuwiki`

2 Organiser Biographies

- **Sébastien Destercke:** He graduated in 2004 as an engineer from the Faculté Polytechnique de Mons in Belgium. In 2008, he earned a Ph.D. degree in computer science from Université Paul Sabatier, in Toulouse (France). He now works as a researcher for the French National Research Center (CNRS), in the Heuristic and Diagnostic of Complex Systems (Heudiasyc) Laboratory, where he leads the AI team as well as the SAFE AI industrial chair. His main research interests are in the field of uncertainty reasoning with imprecision-tolerant models (DS theory, imprecise probabilities, possibility theory, ...), with a focus on issues related to reliability and risk analysis, decision making and machine learning. Sébastien has organized a number of special sessions, workshops and tutorials in various conferences, including ECML, IPMU, ECAI, and has participated in numerous researcher schools as lecturer on the topic of uncertainty reasoning, including SSIPTA, SFLA. . . .

- **Soundouss Messoudi:** She is an Assistant Professor at Université de Technologie de Compiègne (UTC), France. She graduated in 2016 as a computer science engineer from ENSIAS, Morocco. She worked as a BI consultant for a year and a half before obtaining a M.Sc. degree in Machine Learning in 2018 from Paris-Dauphine University, France. In 2022, she completed her PhD at UTC and in the Heuristic and Diagnostic of Complex Systems (Heudiasyc) Laboratory. Her research interests are in the field of trustworthy machine learning. More specifically, she focuses on the development of complex methods allowing for robust and cautious predictions using conformal prediction.

- **Bruce Cyusa Mukama:** He is a doctoral candidate in Computer Science at Université de Technologie de Compiègne (UTC), France. His research, conducted at the Heudiasyc laboratory (UMR-CNRS 7253), centers on developing robust uncertainty quantification algorithms. His most recent contributions are successful applications of conformal prediction in computer vision.

3 Goal of the Proposed Tutorial

Conformal prediction [1] is a rising topic of interest for performing uncertainty quantification (UQ) in machine learning procedures, and more generally to perform UQ from a statistical population.

The goal of the proposed tutorial is :

- to introduce the audience to the basic ideas of conformal prediction, focusing on its key principles, fundamental concepts and ways to apply them in classification and regression problems;
- to provide some details on some recent developments of conformal prediction that could be of interest to the FUZZ'IEEE audience, such as its application to multi-target problems or its use in weakly supervised learning.

While theoretical challenging questions will be mentioned during the tutorial, we will not go deep in this direction, and will mainly give pointers to the more theoretically inclined audience.

4 Plan of the Session

We expect the tutorial to be composed of three main parts, each of them subdivided in smaller parts:

- Part 1: basics of conformal prediction (40 minutes)
  1. Introduction to Conformal Prediction
  2. Conformal Prediction for Classification
  3. Conformal Prediction for Regression
  4. Conformal Prediction for time series (if times allow for it)
- Part 2: conformal predictions for multiple targets (20 minutes)
  1. Multi-variate regression
  2. General multi-target problems
  3. Application to object detection
- Part 3: Conformal prediction for weak labelling (20 minutes)

1. Conformal predictions as weak (fuzzy) labels
2. Learning with weak labels
3. Use in self-training

5 Outline of Covered Material

The tutorial will cover conformal prediction [1] for classification, regression, multi-task learning (i.e., settings that involve classification and regression), and weakly supervised self-training (i.e., settings where the algorithm learns from vague annotations that are not produced by humans). Conformal prediction is an uncertainty quantification framework that can be used to transform machine learning algorithms into set-valued predictors. The sets that are thus produced possess some valuable properties, namely they contain the true value of the predicted variable with a probability level that is chosen by the user, and this is statically guaranteed without necessitating infinitely large data sets nor the knowledge of the data's distribution.

More concretely, when a pet image classifier recognises a "Siberian Husky" in an input that represents an "Alaskan Malamute", for example, conformal prediction can be used to produce a set-valued prediction, such as "{Siberian Husky, Alaskan Malamute}", and to guarantee that this set-valued prediction will contain the actual type of pet with a probability (or confidence level) that is specified by the user. Similarly, when a person has only walked 9500 steps and the pedometer predicts "10k steps", conformal prediction can be used to produce an interval-valued prediction such as "[0.9k steps, 11k steps]". Note that the information in the conformal predictions is more trust-worthy (reliable) even though it is more vague.

Conformal prediction has recently (re)gained popularity as ML algorithms are now being deployed in environments where uncertainty quantification is paramount (safety-critical applications, etc.). This popularity stems from of its simplicity, its efficiency & effectiveness, and its adaptability: it can be used as a post-processing procedure on most prediction algorithms. However, this renewal has also revealed new challenges and opportunities, namely for multi-target prediction problems and our tutorial will cover the state-of-art solutions [3] to these problems as well as some of their practical applications in computer vision [4].

Furthermore, as collecting and annotating large data sets remains a challenge to this day, self-training also remains a prominent technique for the data scientist. Our tutorial will show how conformal prediction can be used to model the uncertainty in the weak (fuzzy) pseudo labels that are used in self-training, and how this can improve the robustness of such learning procedures [2, 5].

6 Justification

- **Audience:** Students and researchers in machine learning, statistics, and anyone interested in understanding and quantifying uncertainty in predictive models.
- **Timeliness:** Conformal prediction has gained tremendous interest in the past years, with both very theoretical contribution (from, e.g., Emmanuel Candes or Robert Tibshirani) to very practical instanciation. Given the natural connection

with fuzzy sets seen as uncertain quantities, we think it would be useful to interest the audience in Fuzz'IEEE to such a framework.

- **Duration:** 90 minutes.
- **Organizer Qualifications:** All organizers have gained significant expertise in conformal prediction, with several publications in the field (S. Destercke has even been a keynote speaker to the conformal flagship conference COPA, and S. Messoudi is the recipient of a best paper award given at COPA). They have both formal and practical contributions to the field, and a very good knowledge of fuzzy sets and related models.

References

[1] Anastasios N Angelopoulos and Stephen Bates. A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *arXiv preprint arXiv:2107.07511*, 2021.

[2] Sébastien Destercke. Uncertain data in learning: challenges and opportunities. *Conformal and Probabilistic Prediction with Applications*, pages 322–332, 2022.

[3] Soundouss Messoudi, Sébastien Destercke, and Sylvain Rousseau. Copula-based conformal prediction for multi-target regression. *Pattern Recognition*, 120:108101, 2021.

[4] Bruce Cyusa Mukama, Soundouss Messoudi, Sylvain Rousseau, and Sébastien Destercke. Copula-based conformal prediction for object detection: a more efficient approach. *Proceedings of Machine Learning Research*, 230:1–18, 2024.

[5] Côme Rodriguez, Vitor Martin Bordini, Sebastien Destercke, and Benjamin Quost. Self learning using venn-abers predictors. In *Conformal and Probabilistic Prediction with Applications*, pages 234–250. PMLR, 2023.