MAPIE LIBRARY : UNIVERSAL UNCERTAINTY QUANTIFICATION IN MACHINE LEARNING

Thibault Cordier[1] | Vincent Blot[1,3] | Louis Lacombe[1] | Thomas Morzadec[1] | Nicolas Brunel[1,2]

1 : Quantmetry, 52, rue d'Anjou, 75008, Paris, France
2 : Laboratoire de Mathématiques et de Modélisation d'Evry, ENSIIE
3 : Laboratoire interdisciplinaire des sciences du numérique, CNRS

Quantmetry Part of Capgemini Invent | MAPIE | 750+ | GitHub

1 MOTIVATION

- Predictions from Machine Learning (ML) models are increasingly important and as more and more use cases being considered at risk, **the need to understand and control the risk** of predictions has become urgent.
- Conformal prediction (CP) is an attractive theoretical framework for **estimating the uncertainties of any predictive algorithms**. Its methodology is general and systematic with few assumptions.
- This framework emphasizes the concepts of readability, transparency, and reliability, hence supporting the principles of **trustworthy AI**.

2 OBJECTIFS

In this work, we contribute to the wide diffusion of the CP framework by developing the MAPIE library that implements such principles:

- Uncertainty quantification for **a variety of ML models and tasks** (regression, classification, CV, NLP, time series) in an unified way.
- **Theoretical guarantees on the coverage** of confidence sets with few hypothesis on the data distribution thanks to conformal predictions.
- Transparency of algorithms with **an open-source implementation**.
- **Industrialization-ready implementation** with tests and documentations.

3 IMPLEMENTING CONFORMAL PREDICTION

- The goal of CP is to **provide a "confidence" set** for each prediction that **ensures a desired marginal coverage** [1]:

For regression: $\mathbb{P}\{Y_{n+1} \in \hat{C}_{n,\alpha}(X_{n+1})\} \geq 1-\alpha$;

For classification: $\mathbb{P}\{Y_{n+1} \in \hat{\mathcal{T}}_{n,\alpha}(X_n+1)\} \geq 1-\alpha$.

With $1-\alpha$: target coverage level; (X_{n+1}, Y_{n+1}) : new test point; $\hat{C}_{n,\alpha}$: prediction interval; $\hat{\mathcal{T}}_{n,\alpha}$: prediction set.

- MAPIE can be **used just like any scikit-learn model** with a fit and a predict method. predict method returns both predictions and prediction intervals/sets.

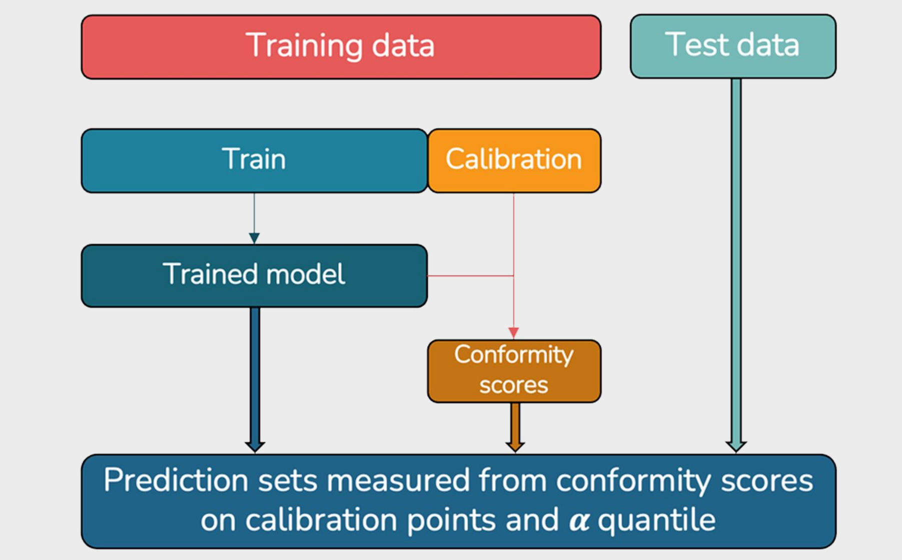

Fig. 1: *Split-Conformal* method involves first training a model and then calibrating it by calculating non-conformity scores to compute the predictions sets. It is an attractive technique due to its computational efficiency (requiring the model to be fitted only once).

Cross-Conformal methods use all data for both training and calibration thanks to resampling methods. They are relevant because they benefit from a compromise between statistical and computational efficiency.

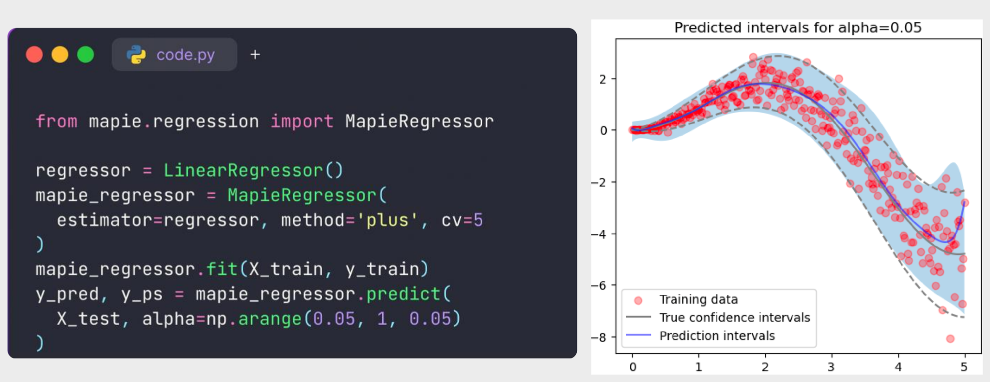

Fig. 2: Regression task - Prediction intervals at risk level $\alpha = 5\%$ with Split-CQR method that allows for a better adaptation of interval widths with heteroscedastic data.

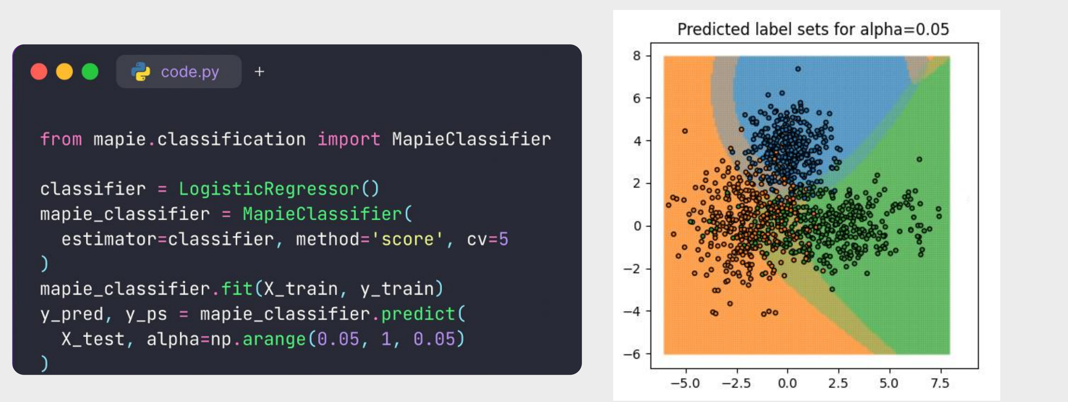

Fig. 3: Classification task - Prediction sets at risk level $\alpha = 5\%$ with LABEL method. Uncertainty is identifiable by a colored overlay area (multi-label prediction set).

Fig. 4: Example of Conformal Prediction with Image Classification on CIFAR10 – MapieClassifier returns prediction sets at some risk level $\alpha = 10\%$. Green the true label of the image if it is included in the prediction set. Red the included labels but which are not the true label. Orange the true label if not included in the prediction set.

- MapieRegressor implements **Split** and **Cross** (**Jackknife+, CV+, Jackknife+-after-Bootstrap**) conformal methods [2,3].
- MapieQuantileRegressor implements **Conformalized Quantile Regression** (CQR) method [4] that allows for a better adaptation with heteroscedastic data.
- MapieClassifier includes **LABEL, Adaptive Prediction Set, Top-K** and **RAPS** methods [5,6,7].
- MapieTimeSeriesRegressor includes **EnbPI** method for dynamic sequential data [8] (non i.i.d. data).

4 MAPIE IN ACTION: CLASSIFICATION / LABELLING

For detection tasks: bring up **the safest detections** to avoid overloading the human operator, or **the most uncertain detection** to direct them to a human operator.
For decision-making tasks: bring up **the safest decisions** to avoid taking an undesired risk, or **the different best decisions** that lead to the desired goal.

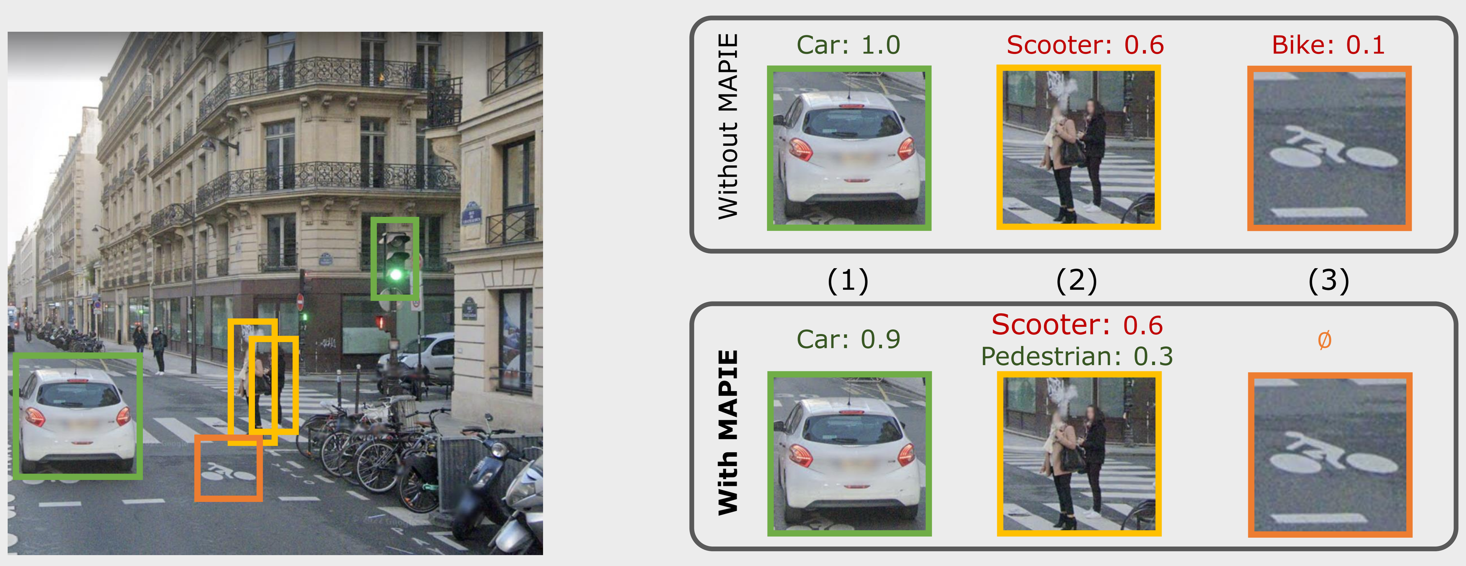

Fig. 5: Illustrative example of how MAPIE approaches CV (Object Detection): (1) It calibrates the scores so that they can be interpreted as probabilities. (2) It can return multiple answers if it is uncertain: the most probable answer might not be correct. (3) It may return nothing because it knows that it does not have enough information to give a pertinent answer.

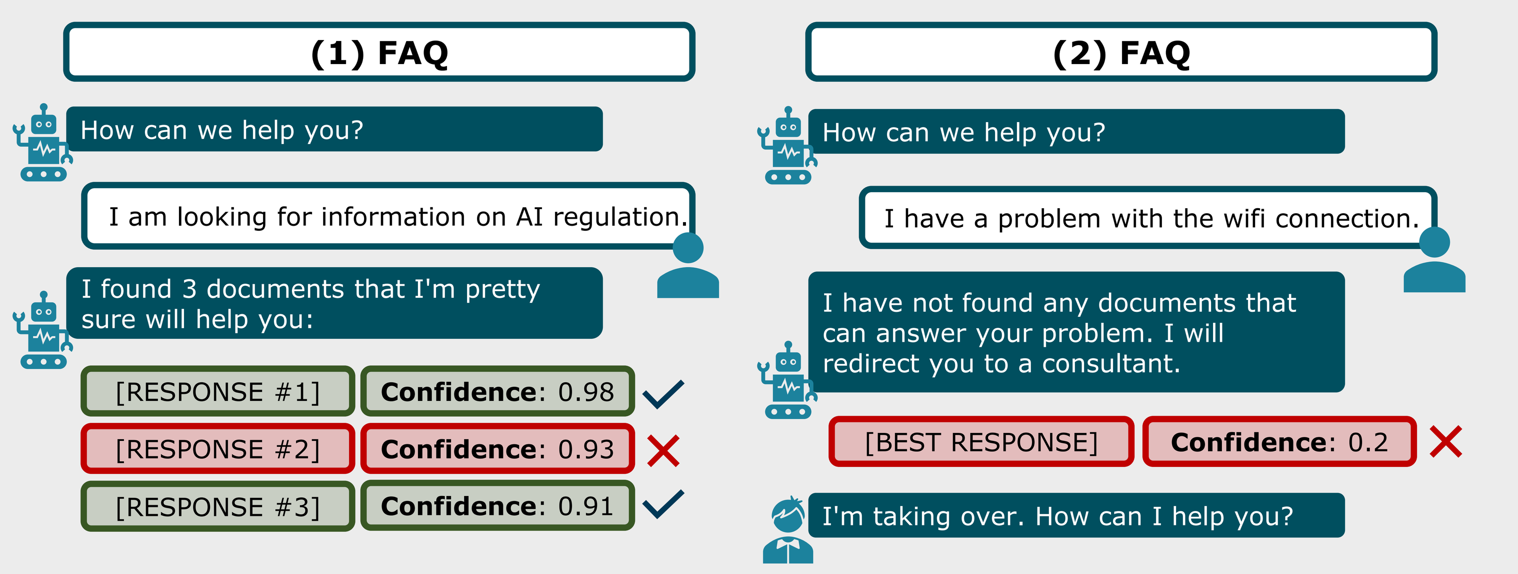

Fig. 6: Illustrative example of how MAPIE approaches NLP (Information Retrieval / Topic Detection / Task-oriented Chatbot):(1) It allows to find the most relevant answers to the user without mental overload by selecting only the most reliable answers. (2) It avoids the frustration of the user if it considers that no answer is pertinent by redirecting him to a human.

5 MAPIE: REGRESSION

For detection tasks: bring up **the safest detections** to avoid returning defective products.

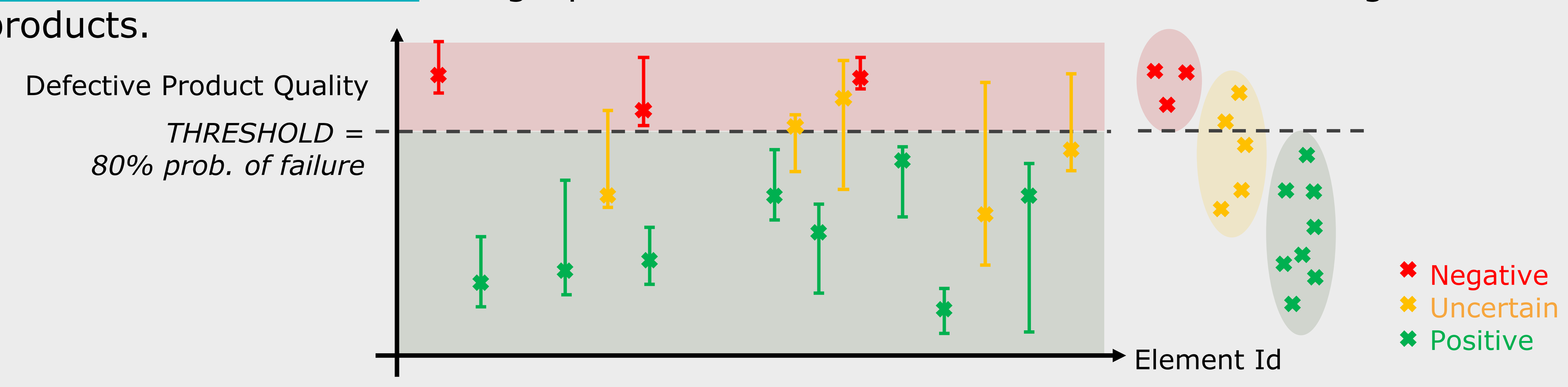

Fig. 7: Illustrative example of how MAPIE approaches regression tasks. To ensure the quality of the products put on the market, we want to predict their longevity during control tests. We keep the positives and reject the negatives. But what to do with those whose quality is uncertain and falls into a risk zone? They can be considered as negative or re-tested depending on the use case.

References :
[1] Anastasios Nikolas Angelopoulos and Stephen Bates. *Conformal Prediction: A Gentle Introduction.* Found. Trends Mach. Learn. 16, 494-591, 2023.
[2] Rina Foygel Barber, Emmanuel J. Candès, Aaditya Ramdas, and Ryan J. Tibshirani. *Predictive inference with the jackknife+*. Ann. Statist., 49(1):486–507, 022021, 2021.
[3] Byol Kim, Chen Xu, Rina Foygel Barber. *Predictive inference is free with the jackknife+-after-bootstrap*. Advances in Neural Information Processing Systems, 33, 4138—4149, 2020.
[4] Yaniv Romano, Evan Patterson, and Emmanuel Candes. *Conformalized quantile regression*. Advances in Neural Information Processing Systems, volume 32. Curran Associates, Inc., 2019.
[5] Mauricio Sadinle, Jing Lei, and Larry Wasserman. *Least Ambiguous Set-Valued Classifiers With Bounded Error Levels*. Journal of the American Statistical Association, 114:525, 223-234, 2019.
[6] Anastasios Angelopoulos, Stephen Bates, Michael Jordan and Jitendra Malik. *Uncertainty Sets for Image Classifiers using Conformal Prediction*. Intl Conference on Learning Representations 2021.
[7] Anastasios Nikolas Angelopoulos, Stephen Bates, Michael Jordan and Jitendra Malik. *Uncertainty Sets for Image Classifiers using Conformal Prediction*. International Conference on Learning Representations 2021.
[8] Chen Xu and Yao Xie. *Conformal Prediction Interval for Dynamic Time-Series*. International Conference on Machine Learning, ICML, 2021.

6 PERSPECTIVES

- A framework for **straightforward implementation**:
  - Modular and flexible architecture, easy-to-use API for prototyping and production purposes.
  - MAPIE also aids in the collaboration with both industrial and academic actors.
- Extension to **new conformal prediction paradigms**:
  - Risk-Controlling Prediction Sets, Learn-Then-Test, Adaptative Conformal Inference, Co-Variate Shift, Multi-target Regression, Binary Classification.
- **They are already using MAPIE:**

Reliable AI | TotalEnergies | RENAULT | MICHELIN | DECATHLON | THALES