Master Internship Offer - Spring 2025
Conformal Prediction from a PAC-Bayesian Perspective

Level: Master 1 or Master 2

Information

Advisor(s): Guillaume Metzler and Stéphane Chrétien
Mail: guillaume.metzler@univ-lyon2.fr stephane.chretien@univ-lyon2.fr
Location *Laboratoire ERIC (Lyon)*

Duration: 4 to 6 months, between March and August 2025
Compensation: 600 euros/month

Perspectives (for M2 student): depending on the outcome of the internship and the candidate's motivation, it may be possible to apply for a thesis grant from the university in order to continue the research in this field.

Keywords: Machine Learning, Statistical Learning, Conformal Prediction, PAC-Bayesian Theory

General Overview: The main idea of this internship is to take advantage of the PAC-Bayesian theoryand study the field Conformal Prediction to study the link between the two and how PAC-Bayesian theory can provide extended guarantees on the prediction made by a model.

Expected profile: Master or engineering degree in Computer Science or Applied Mathematics related to machine/statistical learning. The candidate must show some interest in the theoretical aspects of Machine Learning as well as possess good programming skills. Furthermore, she/he must be fluent in reading and writing in English.

How to apply? Send to guillaume.metzler@univ-lyon2.fr and stephane.chretien@univ-lyon2.fr

- a CV
- your Master grades

Summary

Machine learning algorithms are becoming increasingly widespread in our society. With the rapid expansion of these algorithms, many questions arise concerning their reliability and the generalisation performance when the algorithms under study are applied to new data. For this reason, a deep mathematical analysis of the most widely used algorithms is playing an increasingly important rôle in current research. New tools appear at a fast pace that help investigate why these algorithms generalise well. This research belongs to the main field of **Statistical learning theory** [8]. To derive relevant statistical guarantees, a number of approaches have been developed, such as the notion of uniform stability, the complexity measure of hypothesis spaces or the PAC-Bayesian theory [7, 5]. The PAC-Bayesian approach has recently lead to successive breakthroughs in the discovery of novel generalisation bounds that could not until now be produced using alternative theories. It has also been used to derive new algorithms for minimizing such bounds (known as **self-bounding algorithms** [9, Chapter 7]).

Other approaches have also be used to produce different kinds of statistical guarantees in prediction and risk control, which are various applications to the concept of **Conformal Prediction** [1, 10]. Unlike traditional models that output only a single prediction, conformal prediction assigns a set of possible outcomes associated to each input, ensuring that the true outcome is included in this set with a user-specified probability (e.g., 95%). Hence, **Conformal Prediction** is a statistical framework that provides reliable uncertainty estimates for predictions made by machine learning models. As a particular instance of Conformal Prediction, split Conformal Prediction works by splitting data into a training set and calibration set. The calibration set is used to adjust the prediction sets to statistically guaranteed coverage at the desired confidence level. The strengths of Split Conformal Prediction lies in its inherent simplicity, flexibility, and statistical guarantees. Other approaches have also been developed with narrower prediction sets, such as Jackknife+ [2]. In another direction, the fundamental problem of producing conditional coverage was investigated in recent new works. Indeed, traditional Conformal Prediction techniques usally provide what is called "marginal coverage" guarantees, meaning that the interval length does not depend of the input. Several methods use the quantile regression trick to overcome this issue, thereby ensuring that the prediction intervals contain the true outcome on average across the entire population, with part of the interval beeing conditional of the input. However, these methods may not offer adequate coverage for specific subgroups within the data. Recent results have pushed the limits of Standard Conformal Prediction capabilities by addressing the issue of building schemes that can account for conditional dependency of the prediction interval on the covariates [3].

To sum up, the PAC Bayes and the Conformal based approach both provide generalisation guarantees for the algorithms under study, but from different viewpoints. Let us add that exciting recent work has successfully combined these two tools to derive generalisation guarantees on the coverage properties [6] of an algorithm using conformal prediction, showing that the approaches can be complementary in establishing strong generalisation results.

During the proposed internship, we wish to investigate, in particular, a recent contribution to Conformal Prediction [11], in which the authors manage to enhance the reliability of prediction intervals in Machine Learning models. This approach allows to generate prediction intervals that maintain both marginal coverage and approximate conditional validity for clusters or subgroups naturally present in the data. Our main objectives in this internship will be to: *(i)* reinforce the results obtained in [11] through the lens of PAC-Bayes theory (which applies very well to combinations of models) and *(ii)* investigate exciting applications to **fairness** [4], where notions of subgroups naturally appear when it comes to study different populations.

Main objectives of the proposed Internship During his internship, the candidate will attempt at establishing theoretical guarantees on the predictions made by a model by combining the theories of Conformal Prediction and PAC-Bayesian.

The work can be conducted as follows:

(i) understand the similarities and differences between Conformal Prediction and PAC-Bayesian theory,

(ii) provide additional guarantees on Conformal Prediction using PAC-Bayesian theory,

(iii) study the applicability of the obtained theoretical results to Fairness.

Expected results

- Literature review: PAC Bayesian Learning and Conformal Prediction .
- Theoretical: study how we can improve prediction guarantees provided by the Conformal Prediction theory using the PAC-Bayesian one.
- Practical: Evaluate the derived guarantees and their applicability to Fairness.

For more background material about the topics of this project, the following articles are strongly recommended:

- PAC Bayesian Framework: Pascal Germain *et al.*, *Risk Bounds for the Majority Vote: From a PAC-Bayesian Analysis to a Learning Algorithm*, JMLR, 16(26):787-860, 2015[1]
- Conformal Prediction: *A Gentle Introduction to Conformal Prediction and Distribution-Free Uncertainty Quantification*, Anastasios N. Angelopoulos and Stephen Bates, arXiv 2107-07511, 2022 [2]

[1]https://jmlr.org/papers/v16/germain15a.html.
[2]https://arxiv.org/pdf/2107.07511.

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] Rina Foygel Barber, Emmanuel J Candes, Aaditya Ramdas, and Ryan J Tibshirani. Predictive inference with the jackknife+. 2021.

[3] Isaac Gibbs, John J Cherian, and Emmanuel J Candès. Conformal prediction with conditional guarantees. *arXiv preprint arXiv:2305.12616*, 2023.

[4] Ben Hutchinson and Margaret Mitchell. 50 Years of Test (Un)Fairness: Lessons for Machine Learning. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, page 49–58, 2019.

[5] David McAllester. Some PAC-Bayesian Theorems. In *COLT*, 1998.

[6] Apoorva Sharma, Sushant Veer, Asher Hancock, Heng Yang, Marco Pavone, and Anirudha Majumdar. Pac-bayes generalization certificates for learned inductive conformal prediction. *Advances in Neural Information Processing Systems*, 36, 2024.

[7] John Shawe-Taylor and Robert Williamson. A PAC Analysis of a Bayesian Estimator. In *COLT*, 1997.

[8] Leslie G Valiant. A theory of the learnable. *Communications of the ACM*, 27(11):1134–1142, 1984.

[9] Paul Viallard. *PAC-Bayesian Bounds and Beyond: Self-Bounding Algorithms and New Perspectives on Generalization in Machine Learning*. PhD thesis, Université Jean Monnet de Saint-Etienne, 2023.

[10] Matteo Zecchin, Sangwoo Park, Osvaldo Simeone, and Fredrik Hellström. Generalization and informativeness of conformal prediction. *arXiv preprint arXiv:2401.11810*, 2024.

[11] Yao Zhang and Emmanuel J Candès. Posterior conformal prediction. *arXiv preprint arXiv:2409.19712*, 2024.