Mask-Conditional Conformal Prediction: Valid Uncertainty for All Missing Data Mechanisms

Jiarong Fan[1,2], Juhyun Park[1,3], Thi Phuong Thuy Vo[1,3], Nicolas Brunel[1,3,4]

[1]LaMME, [2]University Paris Saclay, [3]ENSIIE, [4]Capgemini Invent

ABSTRACT

Conformal prediction (CP) offers a principled framework for uncertainty quantification, but it fails to guarantee coverage when faced with missing covariates. In addressing the heterogeneity induced by various missing patterns, Mask-Conditional Valid (MCV) Coverage has emerged as a more desirable property than Marginal Coverage. In this work, we adapt split CP to handle missing values by proposing a preimpute-mask-then-correct framework that can offer valid coverage. We show that our method provides guaranteed Marginal Coverage and Mask-Conditional Validity for general missing data mechanisms. A key component of our approach is a reweighted conformal prediction procedure that corrects the prediction sets after distributional imputation (multiple imputation) of the calibration dataset, making our method compatible with standard imputation pipelines. We derive two algorithms, and we show that they are approximately marginally valid and MCV. We evaluate them on synthetic and real-world datasets. It reduces significantly the width of prediction intervals w.r.t standard MCV methods, while maintaining the target guarantees.

Keywords: Conformal Prediction, Missing Data, Weighted Conformal Prediction, Uncertainty Estimation

Robust Conformal Prediction Using Privileged Information

Shai Feldman[1], Yaniv Romano[1,2]

[1]*Computer Science Department, Technion*

[2]*Electrical and Computer Engineering Department, Technion*

ABSTRACT

We develop a method to generate prediction sets with a guaranteed coverage rate that is robust to corruptions in the training data, such as missing or noisy variables. Our approach builds on conformal prediction, a powerful framework to construct prediction sets that are valid under the i.i.d assumption. Importantly, naively applying conformal prediction does not provide reliable predictions in this setting, due to the distribution shift induced by the corruptions. To account for the distribution shift, we assume access to privileged information (PI). The PI is formulated as additional features that explain the distribution shift, however, they are only available during training and absent at test time. We approach this problem by introducing a novel generalization of weighted conformal prediction and support our method with theoretical coverage guarantees. We further we analyze the robustness of our method to inaccuracies in the weights. Our analysis indicates that our proposal can still yield valid uncertainty estimates even when the weights are poorly estimated. Empirical experiments on both real and synthetic datasets indicate that our approach achieves a valid coverage rate and constructs more informative predictions compared to existing methods, which are not supported by theoretical guarantees.

Keywords: Conformal Prediction, Uncertainty Quantification, Distribution Shift, Corrupted Data, Privileged Information

Adaptive Coverage Policies in Conformal Prediction

Etienne Gauthier[1], Francis Bach[1], Michael I. Jordan[1,2]

[1]*Inria, Ecole Normale Supérieure, PSL Research University*

[2]*Departments of EECS and Statistics, University of California, Berkeley*

ABSTRACT

Traditional conformal prediction methods construct prediction sets such that the true label falls within the set with a user-specified coverage level. However, poorly chosen coverage levels can result in uninformative predictions, either producing overly conservative sets when the coverage level is too high, or empty sets when it is too low. Moreover, the fixed coverage level cannot adapt to the specific characteristics of each individual example, limiting the flexibility and efficiency of these methods. In this work, we leverage recent advances in e-values and post-hoc conformal inference, which allow the use of data-dependent coverage levels while maintaining valid statistical guarantees. We propose to optimize an adaptive coverage policy by training a neural network using a leave-one-out procedure on the calibration set, allowing the coverage level and the resulting prediction set size to vary with the difficulty of each individual example. We support our approach with theoretical coverage guarantees and demonstrate its practical benefits through a series of experiments.

Keywords: Conformal prediction; E-values; Uncertainty quantification; Machine learning

Towards a Rigorous Evaluation of RAG Systems: The Challenge of Due Diligence

Grégoire Martinon, Alexandra Lorenzo de Brionne, Jérôme Bohard, Antoine Lojou, Damien Hervault, **Nicolas Brunel**

Invent Lab, Capgemini Invent, France and ENSIIE, LaMME, Université Paris Saclay

ABSTRACT

The rise of generative AI has driven significant advancements in high-risk sectors like healthcare and finance. The Retrieval-Augmented Generation (RAG) architecture, combining language models (LLMs) with search engines, is particularly notable for its ability to generate responses from document corpora. Despite its potential, the reliability of RAG systems in critical contexts remains a concern, with issues such as hallucinations persisting. This study evaluates a RAG system used in due diligence for an investment fund. We propose a robust evaluation protocol combining human annotations and LLM-Judge annotations to identify system failures, like hallucinations, off-topic, failed citations, and abstentions. Inspired by the Prediction Powered Inference (PPI) method, we achieve precise performance measurements with statistical guarantees. We provide a comprehensive dataset for further analysis. Our contributions aim to enhance the reliability and scalability of RAG systems evaluation protocols in industrial applications. To ensure consistency in the online technical program, this abstract template must be used for submissions.

Keywords: LLM, RAG, hallucinations, LLM-as-a-Judge