Conformal Prediction: Calibrated Decision-Making

Akash Singh
HEC - University of Liege
email:Akash.singh@uliege.be

Ashwin Ittoo
HEC - University of Liege
email:Ashwin.ittoo@uliege.be

Pierre Ars
Lead Actuarial Innovation, Ethias
email:Pierre.ars@ethias.be

Elise Vandomme
HEC - University of Liege
email:Elise.vandomme@uliege.be

Keywords: Uncertainty, Conformal prediction, Machine learning

Quantifying uncertainty is crucial for assessing the trustworthiness of Machine learning (ML) predictions. Factors like data noise or transformations can influence uncertainty in ML predictions. Standard ML models like Random Forest(RF), XGBoost (XGB), and Neural Networks(NN) provide accurate point predictions but lack calibrated confidence or statistical guarantees, often producing uncalibrated probabilities without inherent uncertainty representation. While these methods perform well, they are not inherently designed to provide coverage guarantees. Coverage guarantee is theoretical assurance that the prediction sets produced by the conformal method will contain the true label of a given instance with a specified probability $1 - \alpha$, where α is the chosen significance level [1]. Conformal prediction (CP) [1, 2] addresses these limitations by providing explicit coverage guarantees through prediction intervals or calibration sets, correcting the miscalibrated probabilities, and ensuring that the true value lies within the interval with a specified probability (e.g., 90%). A CP set is a set of all possible class labels that a given data point could be classified into. Additionally, CP is a model-agnostic framework that is flexible and versatile and can be applied to any ML model.

In this work, we apply CP to a leading Belgian insurance company's data and binary churn prediction model and to several widely available public churn and fraud prediction datasets from prior research [4, 5, 6]. In a binary churn classification model, each customer's profile x is assigned to one of two discrete classes i.e. *churn (0)* or *not churn(1)*. For a binary classification problem, the "CP set" can take one of the following four forms: $\emptyset, \{0\}, \{1\}, \{0, 1\}$. Conformal Prediction generalizes the concept of hypothesis testing by calculating *p-values* for each potential class. We further define the Significance level (α) that specifies the desired coverage level ($1 - \alpha$). For instance, if α=0.1, the prediction set is designed to contain the true label at least 90% of the time. The threshold in conformal prediction is a value that determines the size of the prediction set. It's calculated based on the distribution of nonconformity scores. Thus if the *p-value* for a class exceeds a threshold (based on the "significance level α"), that label is included in the prediction set. The threshold is derived from the α-quantile of a

set of nonconformity scores.
To illustrate the methodology, we represent a profile in the portfolio with x where **significance level** is $\alpha = 0.1$ with corresponding **threshold** of 0.30. Let $P(C|x)$ denote the churn probability output of a predictive model (e.g., RF or XGB). Let $CC(x)$ represent the "Conformal Prediction set (or class)".
The crucial result is that it can be proven that the final decision regarding the profile x (churn or not churn) belongs to the Conformal Prediction set $CC(x)$ with a probability of at least $1 - \alpha$ [1]. This result applies globally to the entire dataset for both classes.
For example consider three instances of profile x_1, x_2, x_3 where each instance yields a distinct CP set.
1. **Profile** x_1: Assume $P(C|x_1) = 0.15$. Then $P(C|x_1) < 0.30$, we have $1 - P(C|x_1) \geq 0.30$. Therefore, $CC(x_1) = \{0\}$. With significance level $\alpha = 0.1$, the conformal prediction for x_1 not churning is: $P(x_1 \text{ will not churn}) \geq 0.90$.
2. **Profile** x_2: Assume $P(C|x_2) = 0.75$. So $P(C|x_2) \geq 0.30$ and $1 - P(C|x_2) \leq 0.30$. Therefore, $CC(x_2) = \{1\}$. Because $\alpha = 0.1$ the conformal prediction of x_2 will churn is: $P(x_2 \text{ will churn}) \geq 0.90$.
3. **Profile** x_3: Assume $P(C|x_3) = 0.60$. Then $P(C|x_3) \geq 0.30$ and $1 - P(C|x_3) \geq 0.30$. Thus $CC(x_3) = \{0, 1\}$. For x_3, there is insufficient evidence to confidently classify it as either churn or not churn with a confidence level of at least 0.90. The Conformal Prediction set $\emptyset$ can occur if, for example, the threshold is 0.60 and $P(C|x) = 0.55$. The CP framework enables decision-makers to reduce uncertainty. For example, a fraud manager can confidently act on cases classified as solely fraudulent or non-fraudulent (CP sets 0 or 1), with confidence, knowing the fraud probability is less than 10% or at least 90%, respectively. Uncertain cases (CP set {0,1}) can be flagged for further analysis. CP offers reliable uncertainty quantification, calibrated probabilities, flexible decision-making, and theoretical guarantees, addressing the limitations of traditional machine learning.

References

[1] Vladimir Vovk, Alexander Gammerman, and Glenn Shafer, *Algorithmic Learning in a Random World, Second Edition*, Volume 29, Springer, 2022.

[2] Eyke Hüllermeier and Willem Waegeman, *Aleatoric and Epistemic Uncertainty in Machine Learning: An Introduction to Concepts and Methods*, Machine Learning, Volume 110, Number 3, Pages 457–506, Springer, 2021.

[3] Ryan Tibshirani, *Conformal Prediction, Advanced Topics in Statistical Learning*, Spring 2023, Berkeley.

[4] Aman Chauhan, *Fraud Detection*, Kaggle.com, Available at: `https://www.kaggle.com/datasets/whenamancodes/fraud-detection/data`, 2015.

[5] BlastChar, *Telco Customer Churn*, Available at: `https://www.kaggle.com/datasets/blastchar/telco-customer-churn`, 2018.

[6] Shubham Kumar, *Churn Modelling Bank*, Available at: `https://www.kaggle.com/datasets/shubh0799/churn-modelling`, 2020.