Beyond Yes or No:

Making Reliable Decisions and improving personalized customer service

Introduction

- Noise in data confuses the model, hindering its ability to distinguish true patterns from random. Thus, we need to know when it's just guessing (i.e. low confidence).
- When regular machine learning gives a prediction, it does not tell you how confident it is, and that missing "confidence information" is important in situations where errors could be dangerous.
- New regulations (e.g. EU AI Act) emphasize trustworthy, unbiased models.
- With customer churn in insurance as a case study, we show how knowing a model's confidence can improve customer service.

Methods

1. **HEUQ** (Heterogeneous Ensemble for Uncertainty Quantification)
   - Quantifies uncertainty by assessing disagreement among models.
   - Combines 6 diverse models to estimate
     - Total Uncertainty : Model + Data
     - Epistemic Uncertainty: Model
     - Aleatoric uncertainty: Data/Noise

   $U_t(x) = U_e(x) + U_a(x)$

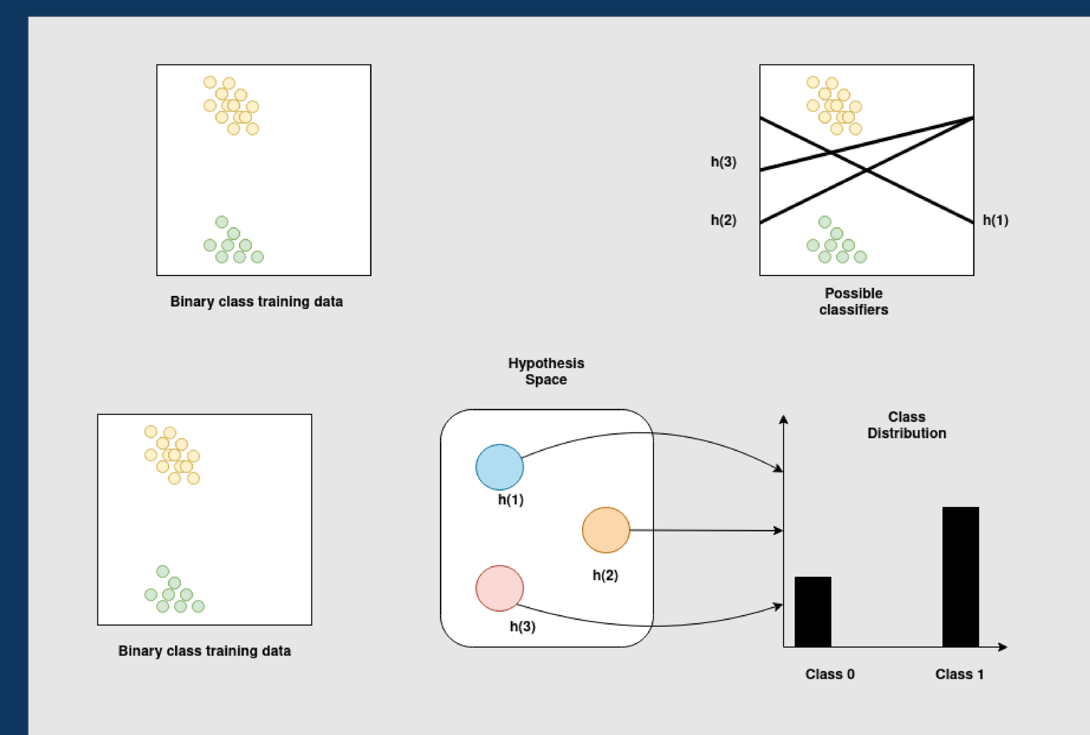
2. **CP** (Conformal Prediction)
   - Transforms raw predictions into statistically sound ones with a predefined confidence level
   - Provides calibrated uncertainty sets

3. Sub Portfolio Selection based on **UQ** to optimize accuracy

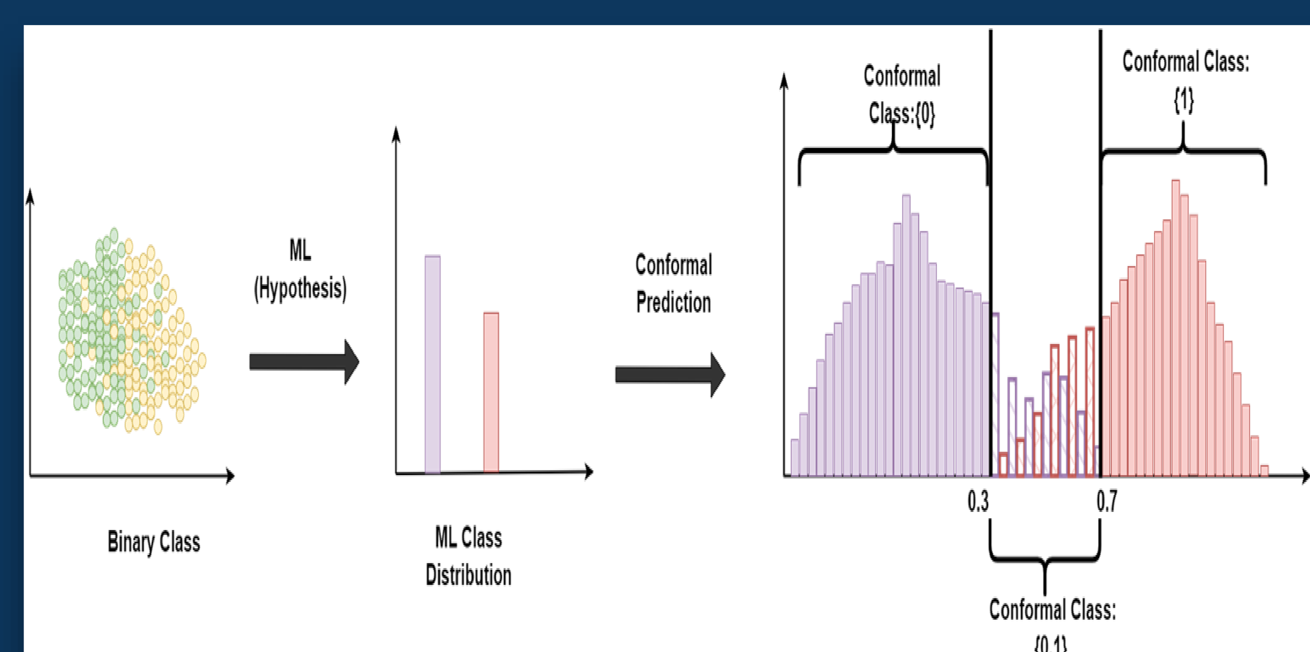
HEC LIÈGE Management School - Liège Université

ethias

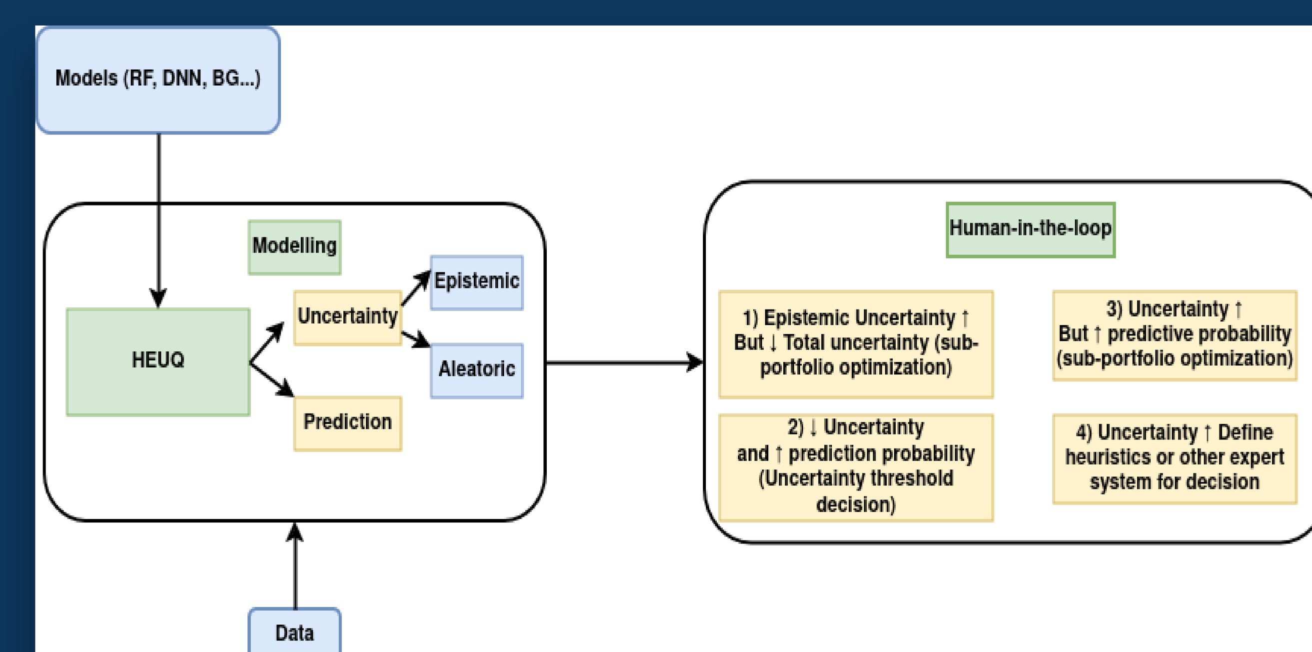
Uncertainty: From liability to a Valuable asset for improved decision making



An e.g. of binary classification. Multiple optimal hypothesis can exist in hypothesis space that can be used to classify. This Fig. shows how there can be multiple optimal hypothesis learned differently .



This figure illustrates binary Conformal Prediction. With a 0.30 threshold, we predict Class 0 or Class 1 with 90% certainty. When confidence is low, the prediction is a set of {Class 0, Class 1}, explicitly showing model uncertainty and preventing overconfident decisions



This figure demonstrates how an organization can use uncertainty information from machine learning models to effectively bring humans into the decision-making loop. This is especially valuable for situations where the AI model's predictions are under-confident. By highlighting these uncertain cases, the system can flag them for human review, ensuring more reliable and informed decisions, particularly in critical applications.

Results

- Targeted sub-portfolios selected UQ achieved up to 74.34% balanced accuracy.
- Conformal prediction at 4% confidence identified loyal customers, with churn rates as low as 2.5%.

U_e increasing →	Balanced Accuracy							
				80.01	78.58	82.51	84.64	79
U_t increasing ↓	63.89	74.35	69.24	69.72	70.84	73.24	71.55	70.3
	60.25	68.08	64.41	67.73	67.41	65.45	65.39	65.72
	58.42	63	62.64	58.92	62.54	63.2	64.59	58.38
	55.27	57.26	57	57.6	55.92	60.71	53.72	56.78
	53.15	52.66	50.45	53.32	52.5	53.22	49.18	52.31

Balanced accuracy (BA) for sub-portfolios obtained by crossing of the 8-quantiles for Ue and 6-quantiles for U_t. For example, 62.54 is the BA for the sub-portfolio composed by profiles satisfying $u_e(x)$ being between 4th 8-quantile and 5th 8-quantile and $U_t(x)$ being between 3th 5-quantile and 4th 5-quantile. The marketing manager can then select sub portfolios based on $U_e(x)$ an $U_t(x)$ according to his objectives in terms of BA

Applications

- Identify optimal sub portfolios with higher precision scores
- Improve balanced accuracy by selecting customers based on uncertainty levels
- Enable human-in-the-loop decision making for high-uncertainty cases

Discussion

- Enhanced decision-making with a nuanced understanding of prediction confidence
- Enabling more personalized customer service strategies
- Improving risk management through a better understanding of prediction reliability
- Building trust in AI systems through confidence levels

Scan to download the full paper and Poster