Symbolize Regressions with PHAL

Fraunhofer IIS

Post-Hoc model Approximation with Logic

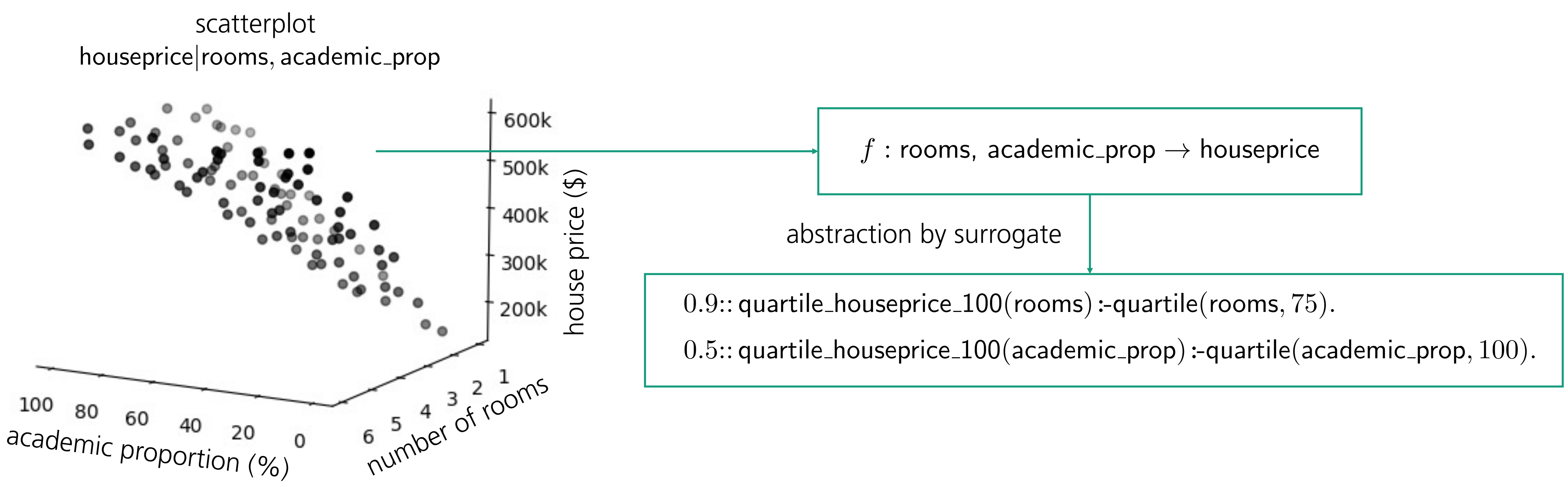
Emanuel Slany
Research funded by German Federal Ministry of Education and Research, Project Human-Centered AI in Chemical Industry

Motivation

- **Problem**: Regression models are generally built upon domain-specific expert knowledge. Their ex-post extension by domain experts with novel knowledge is often infeasible due to the models' complexity.
- **Solution**: We symbolize regression models with probabilistic logic. The resulting surrogate model actively considers user preferences.
- **Contribution**: We reduce the barriers to induce knowledge into regression models and show that our method is superior to the state of the art.

Approach

- **Goal**: Find a probabilistic logical surrogate model.
- **Abstraction step**: Reduce the level of measurement of features and regression target by user-defined statistical feature extraction procedures.
- **Rule learning step**: **1)** Convert extracted features into probabilistic logical predicates. **2)** Construct examples consisting out of abstracted features and target pairs. **3)** Apply the rule learner ProbFOIL+.



Example 1. Consider `Charlestown` (Boston) as a fictive example with an average house price of $\$750,000$ and an average of 4 rooms per house. Let us say that in `Charlestown`, the house price is caused by the room number with a probability of 90%, which can be expressed in probabilistic logic as:
0.9 :: houseprice(`Charlestown`, $\$750,000$) :- rooms(`Charlestown`, 4).

Example 2. Assume that we cannot capture whether an average of 4 rooms per house is high or low. We abstract the scalar 4 into its corresponding quartile value by: quartile(rooms(`Charlestown`)) $= 75$, which means that 75% of all districts in Boston on average have 4 or less rooms per house.

Definition 1 (Inverse Coefficient of Variation)**.** Let the *inverse coefficient of variation* (inverse cv) approximate the probability of y to be true. Then, $y_{prob} = 1 - cv$, where $cv = y_{cov} \times (y_{mean} + l)^{-1}$ with $l > 0$ and $cv = [0, 1]$.

Example 3. We set $l = 0.01$ to account for divisions by 0. We know from Ex. 1 the average house price of `Charlestown`, which is $\$750,000$. Assume for now that it origins in a GPR with $y_{cov} = \$375,000$. Inserting the numbers into the equation of Def. 1 yields: $y_{prob} = 1 - (375,000 \times (750,000 + 0.01)^{-1}) \approx 0.5$.

Example 4. Let us translate the average room number of `Charlestown`, the only entry in $\mathcal{X}_{Exp}$, into ProbLog syntax: $t(x_1^{(1)})$ = p_1_1. The feature name is retrieved by: $name(x_1)$ = rooms. Referencing Ex. 2, we straightforward formulate: quartile(p_1_1, 75). We specify the target predicate quartile_houseprice_100 and assume that it is satisfied by `Charlestown`. Recycling y_{prob} from Ex. 3 gives:
0.5 :: quartile_houseprice_100(p_1_1).

Evaluation

Table 1: Surrogate quality. We highlight Welch's tests with $p \leq 0.05$ (m^*).

Data	Method	Δ-fp	Δ-fn	# Rules	∅ Features	S
Housing	PHAL	**0.04** (0.03)	**0.11*** (0.09)	7.20 (0.84)	2.87 (0.51)	**0.85*** (0.04)
Housing	GridEx	0.09 (0.05)	0.39 (0.40)	**5.00*** (0.00)	**1.40*** (0.55)	0.19 (0.01)
Wine	PHAL	**0.04** (0.02)	**0.07** (0.04)	**2.80*** (0.84)	3.68 (0.29)	**0.87*** (0.06)
Wine	GridEx	0.09 (0.05)	0.47 (0.20)	4.40 (1.14)	**1.40*** (0.89)	0.26 (0.03)

- **Strategy**: Comparison of PHAL to GridEx for Housing and Wine data sets wrt. fidelity (Δ-fp , Δ-fn), complexity (# Rules , ∅ Features), and stability (S).
- **Results**: Whereas the surrogate's complexity does not reveal a clear trend, PHAL is superior in terms of fidelity and stability.

Outlook

- Enrich PHAL with the ability to return continuous values.
- Conduct further experiments with several feature extraction procedures.
- Apply PHAL to parameterization in the automotive sector.
- Exploit PHAL to develop explainable and interactive training procedures for regression models.

1 Luc De Raedt and Angelika Kimmig. Probabilistic (logic) programming concepts. Mach. Learn., 100(1):5-47, 2015.

2 Luc De Raedt, Anton Dries, Ingo Thon, Guy Van den Broeck, and Mathias Verbeke. Inducing Probabilistic Relational Rules from Probabilistic Examples. In Qiang Yang and Michael J. Wooldridge, editors, Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015, pages 1835-1843. AAAI Press, 2015.

3 Federico Sabbatini, Giovanni Ciatto, and Andrea Omicini. GridEx: An Algorithm for Knowledge Extraction from Black-Box Regressors. In Davide Calvaresi, Amro Najjar, Michael Winiko , and Kary Främling, editors, Explainable and Transparent AI and Multi-Agent Systems - Third International Workshop, EXTRAAMAS 2021, Virtual Event, May 3-7, 2021, Revised Selected Papers, volume 12688 of Lecture Notes in Computer Science, pages 18-38. Springer, 2021.

4 Christian Wirth, Ute Schmid, and Stefan Voget. Humanzentrierte Künstliche Intelligenz: Erklärendes interaktives maschinelles Lernen für Effizienzsteigerung von Parametrieraufgaben. In Ernst A. Hartmann, editor, Digitalisierung souverän gestalten II, pages 80-92, Berlin, Heidelberg, 2022. Springer Berlin Heidelberg.

Nordic Probabilistic AI (ProbAI) Summer School, Trondheim, June 12-16, 2023.

Fraunhofer IIS
Project Group Comprehensible AI
Federal Ministry of Education and Research