Causal Reasoning with Probabilistic Graphical Models

Probabilistic Graphical Models - Project Proposal - 5th November 2018

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Moritz Willig 2527455; Johannes Czech 2422372

1 Motivation

The aspect of causal reasoning is raising interest in today's research (Doshi-Velez and Kim (2017)). In 2018 the European Commission attaches great importance to the interpretability of artificial intelligence and AI-safety (*Interpretability in AI and its relation to fairness, transparency, reliability and trust - JRC Science Hub Communities - European Commission* (2018). Probabilistic graphical models (PGM) have proved to be an effective method for expressing conditional dependencies among a set of random variables. The application for PGM in the context for causal reasoning has been demonstrated by Pearl (2013) as well as Eichler and Didelez (2007)

2 Goals

Our goal is to evaluate the usage of PGMs on the *Census Is Not Adult* dataset (Guyon, Aliferis, Cooper and Elisseeff (2007)). The CINA-dataset has been presented in the *Challenge: Datasets - Causality Workbench* (2007) with the task to detect the socio-economic factors which encourage high income. The overall dataset consists of 16,033 train samples and 10,000 test samples with 132 features. These features include categorical as well as numeric variables such as age, workclass, education, education, marital status, occupation and native country. Additionally artificially generated distracting features were added that aren't a cause of the target variable. The target value is binary and indicates whether the annual income of a person exceeds $50K. Originally this dataset is derived from the census data provided by the UCI machine-learning repository Adult database (US Census Bureau Applications Development {and} Services Division (2007)).

3 Evaluation

In order to determine a winner the committee used the the test score for evaluating the target prediction values. Moreover they propose the following metrics for a more detailed view on the model predictions (*Challenge: Evaluation - Causality Workbench* (2007)):

Causal Discovery

- **Fnum:** The number of features or the best number of features used to make predictions with.
- **Fscore:** Score for the list of features provided. For sorted list, the most predictive features should come first to get a high score. For unsorted lists, the features provided should be highly predictive to get a high score.

Target Prediction

- **Dscore:** Discovery score evaluating the target prediction values.
- **Tscore:** Test score evaluating the target prediction values.

We will evaluate our probabilistic model using these metrics and compare our results with the leaderboard table.

References

Challenge: Datasets - causality workbench. (2007). Retrieved 2018-11-01, from http://www.causality.inf.ethz.ch/challenge.php?page=datasets

Challenge: Evaluation - causality workbench. (2007). Retrieved 2018-11-05, from http://www.causality.inf.ethz.ch/challenge.php?page=evaluation#cont

Doshi-Velez, F. & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. Retrieved 2018-11-01, from http://arxiv.org/abs/1702.08608

Eichler, M. & Didelez, V. (2007). Causal reasoning in graphical time series models. , 8.

Guyon, I., Aliferis, C., Cooper, G. & Elisseeff, A. (2007). Datasets of the causation and prediction challenge. , 89.

Interpretability in AI and its relation to fairness, transparency, reliability and trust - JRC science hub communities - european commission. (2018). Retrieved 2018-11-01, from https://ec.europa.eu/jrc/communities/community/humaint/article/interpretability-ai-and-its-relation-fairness-transparency-reliability-and

Pearl, J. (2013). Graphical models for probabilistic and causal reasoning. , 29.

US Census Bureau Applications Development {and} Services Division, W. (2007). *US census bureau: Page not found.* Retrieved 2018-11-05, from https://www.census.gov/DES/www/welcome.html