Modeling Robustness in Decision-Focused Learning as a Stackelberg Game

Extended Abstract

Sonja Johnson-Yu
Harvard University
sjohnsonyu@g.harvard.edu

Kai Wang
Harvard University, Google Research
kaiwang@g.harvard.edu

Jessie Finocchiaro
Center for Research on Computation and Society, Harvard University
jessie@seas.harvard.edu

Aparna Taneja
Google Research
aparnataneja@google.com

Milind Tambe
Harvard University, Google Research
tambe@g.harvard.edu

ABSTRACT

Predict-then-optimize is a common paradigm for optimization tasks situated in incomplete informational settings, in which an agent estimates missing parameters and then optimizes over these predicted parameters. One proposed improvement to this predict-then-optimize framework is *decision-focused learning*, which establishes an end-to-end learning pipeline, allowing a predictive model to be tailored to the particular optimization task. The behavior of this predict-then-optimize framework in the presence of noise, however, is not well-understood. This is problematic because many data collection and annotation systems are inherently noisy, and the introduction of such noise could lead to poor downstream optimization. In this work, we aim to present results on robustness to label noise in decision-focused learning and traditional predict-then-optimize tasks using a Stackelberg game as the underlying framework of explanation. Our results suggest that playing the Stackelberg game in anticipation of label noise yields robustness in the predict-then-optimize framework at large, and that the optimal decision-focused learning Stackelberg solution continues to outperform the optimal traditional predict-then-optimize Stackelberg solution.

KEYWORDS

Robust Optimization; Predict-Then-Optimize; Adversarial Attack and Defense

ACM Reference Format:
Sonja Johnson-Yu, Kai Wang, Jessie Finocchiaro, Aparna Taneja, and Milind Tambe. 2023. Modeling Robustness in Decision-Focused Learning as a Stackelberg Game: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 2 pages.

1 INTRODUCTION

Autonomous agents often incorporate predictive AI models to make "smart" decisions in the midst of incomplete information using the *predict-then-optimize* framework [4, 7, 8, 10, 13]. In predict-then-optimize, an agent predicts missing information, and then optimizes a reward function based on these predictions. The standard way of accomplishing this is to train a *two-stage* model in which the agent first learns a machine learning model trained to maximize predictive accuracy, and then runs an optimization algorithm maximizing a *decision quality function* over the trained model's predictions. Notably, maximizing predictive accuracy is not always equivalent to maximizing the specified reward function [6]. To address this issue, *decision-focused learning*, in contrast to two-stage learning, trains the predictive model to optimize the decision quality function and differentiates through the entire prediction and optimization pipeline, making training an end-to-end process. Decision-focused learning has been shown to improve performance over traditional two-stage models across a variety of domains and applications such as traffic navigation optimization, portfolio optimization, web advertising, and allocating scarce maternal health resources [6, 8, 11, 13].

While decision-focused learning is becoming increasingly popular, little is known about its performance in "messy" settings, like those with noisy or sparse data. Three possible sources of label noise include: (a) sparsity in the training data, (b) bias in training data collection, and (c) distribution shift between the training and test sets. While decision-focused learning has historically outperformed standard two-stage models, we have no understanding if this holds in the presence of noisy labels at test time.

Butler et al. [5] demonstrate that both two-stage and decision-focused learning are susceptible to poisoning attacks that add adversarial noise to the features that the model is trained on [12]. However, to the best of our knowledge, no work has compared the relative robustness of two-stage and decision-focused learning, nor have any works proposed methods for improving the robustness of decision-focused learning in the presence of noisy labels.

In this work, we cast two-stage and decision-focused learning under label noise as general-sum and zero-sum Stackelberg games, respectively, and use this modeling to derive bounds on the relative performance of the two frameworks. Given the growing popularity of algorithmic agents that make decisions via the predict-then-optimize framework and the frequency of noisy data labels, this work is the first to our knowledge that explores the intersection between predict-then-optimize and robust game theory (cf. [1, 9]).

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

2 BACKGROUND

2.1 The Learner's Predict-then-optimize Problem

In the standard predict-then-optimize framework, an autonomous agent predicts something about the state of the world, then optimizes a decision quality function given their prediction. The predictive task is to learn a parameterized model m_w from given features $x \in \mathcal{X}$ to predict the unknown parameters $y \in \mathcal{Y} \subseteq \mathbf{R}^k$. The optimization problem is to then maximize a *decision quality function* $f : \mathcal{Z} \times \mathcal{Y} \to \mathbf{R}$ as a function of decision $z \in \mathcal{Z}$ and parameters $y \in \mathcal{Y}$.

In this framework, we assume the learner first makes a prediction $\hat{y} \coloneqq m_w(x)$ using the predictive model $m_w : \mathcal{X} \to \mathcal{Y}$ parameterized by weights w. The learner then uses the prediction $\hat{y}$ as the parameter of the optimization problem to find the optimal decision $z^*(\hat{y})$, where z^* is defined by

$$z^*(\hat{y}) \coloneqq \arg\max_{z \in \mathcal{Z}} f(z, \hat{y}) \ . \tag{1}$$

The decision is then evaluated on the ground truth parameter y to obtain a decision quality $f(z^*(\hat{y}), y)$.

The learner is given a dataset $\mathcal{D}_{\text{train}} = \{x_i, y_i\}$ to train the predictive model. After the model m is trained, a testing dataset $\mathcal{D}_{\text{test}}$ is presented. The learner uses the given features to generate predictions of the missing labels and propose the corresponding decisions. The decisions are evaluated on the revealed ground truth labels in the testing set:

$$\frac{1}{|\mathcal{D}_{\text{test}}|} \sum_{(x,y) \in \mathcal{D}_{\text{test}}} f(z^*(\hat{y}), y), \quad \hat{y} = m_w(x)$$

We are primarily concerned with the setting in which the labels $y_i \in \mathcal{D}_{\text{test}}$ are noisy.

2.2 Learning Methods without Label Noise

We summarize two existing learning methods with different objectives that the learner uses to train the predictive model m_w parameterized by the weight w.

The two-stage (TS) approach learns a predictive model m_w by minimizing mean squared error:

$$\min_w \sum_{(x,y) \in \mathcal{D}_{\text{train}}} \|m_w(x) - y\|_2 \tag{TS}$$

After the model is learned, the predictions $\hat{y} = m_w(x)$ are then used to optimize the decision quality function $z^*(\hat{y})$ in Equation 1.

In contrast, the predictive objective in decision-focused learning is the decision quality function instead of mean squared error.

$$\max_w \sum_{(x,y) \in \mathcal{D}_{\text{train}}} f(z^*(m_w(x)), y) \tag{DFL}$$

The advantage of decision-focused learning is the alignment of the training objective and the testing objective. To optimize the objective in DFL, it is common to use gradient descent, which requires back-propagating through the optimal decision z^* defined in Equation 1 . This can be achieved by differentiating through the optimality and KKT conditions as shown by [2, 3].

3 CONTRIBUTION

In this paper, we propose to study the learning challenge in a predict-then-optimize framework with a potential mismatch in the training distribution and the testing distribution. We cast the learning problem as Stackelberg games; these formulations lead to robust two-stage and robust decision-focused learning problems. We show bounds on the performance of the robust two-stage and the robust decision-focused learning problems with a tightness guarantee.

4 ACKNOWLEDGMENTS

Research was sponsored by the ARO and was accomplished under Grant Number: W911NF-18-1-0208. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of ARO or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. Additionally, this material is based upon work supported by the National Science Foundation under Award No. 2202898.

REFERENCES

[1] Michele Aghassi and Dimitris Bertsimas. Robust game theory. *Mathematical programming*, 107(1):231–273, 2006.
[2] Akshay Agrawal, Brandon Amos, Shane Barratt, Stephen Boyd, Steven Diamond, and Zico Kolter. Differentiable Convex Optimization Layers. *arXiv:1910.12430 [cs, math, stat]*, October 2019. URL http://arxiv.org/abs/1910.12430. arXiv: 1910.12430.
[3] Brandon Amos and J. Zico Kolter. OptNet: Differentiable Optimization as a Layer in Neural Networks. In *Proceedings of the 34th International Conference on Machine Learning*, pages 136–145. PMLR, July 2017. URL https://proceedings.mlr.press/v70/amos17a.html. ISSN: 2640-3498.
[4] Mallik Angalakudati, Siddharth Balwani, Jorge Calzada, Bikram Chatterjee, Georgia Perakis, Nicolas Raad, and Joline Uichanco. Business analytics for flexible resource allocation under random emergencies. *Management Science*, 60(6): 1552–1573, 2014.
[5] Ryan Butler, Wong Wai Tuck, Anuresh Sinha, and Thanh Ngyuen. Poisoning attacks on data-based decision making: A preliminary study. *AASG-22: 3rd Autonomous Agents for Social Good (AASG) held at the 21st International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, May 2022.
[6] Chris Cameron, Jason Hartford, Taylor Lundy, and Kevin Leyton-Brown. The perils of learning before optimizing. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(4):3708–3715, Jun. 2022. doi: 10.1609/aaai.v36i4.20284. URL https://ojs.aaai.org/index.php/AAAI/article/view/20284.
[7] Carri W Chan, Vivek F Farias, Nicholas Bambos, and Gabriel J Escobar. Optimizing intensive care unit discharge decisions with patient readmissions. *Operations research*, 60(6):1323–1341, 2012.
[8] Adam N. Elmachtoub and Paul Grigas. Smart "Predict, then Optimize". *Management Science*, 68(1):9–26, January 2022. ISSN 0025-1909. doi: 10.1287/mnsc.2020.3922. URL https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2020.3922. Publisher: INFORMS.
[9] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples, 2014. URL https://arxiv.org/abs/1412.6572.
[10] Mili Mehrotra, Milind Dawande, Srinagesh Gavirneni, Mehmet Demirci, and Sridhar Tayur. Or practice—production planning with patterns: A problem from processed food manufacturing. *Operations research*, 59(2):267–282, 2011.
[11] Sanket Shah, Bryan Wilder, Andrew Perrault, and Milind Tambe. Learning (local) surrogate loss functions for predict-then-optimize problems. *Neural Information Processing Systems*, 2022.
[12] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. Intriguing properties of neural networks. *arXiv preprint arXiv:1312.6199*, 2013.
[13] Bryan Wilder, Bistra Dilkina, and Milind Tambe. Melding the Data-Decisions Pipeline: Decision-Focused Learning for Combinatorial Optimization. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):1658–1665, July 2019. ISSN 2374-3468. doi: 10.1609/aaai.v33i01.33011658. URL https://ojs.aaai.org/index.php/AAAI/article/view/3982. Number: 01.