Interactions Between Learning and Decision Making

Theja Tulabandhula (Massachusetts Institute of Technology; theja@alum.mit.edu)
DOI: 10.1145/3054837.3054847

This thesis revolves around learning and decision making and how these two processes interact with each other in surprising ways. To this end, we have studied the effects of machine learning on decision making and vice versa in a variety of theoretical contexts. On the way, we have also touched application areas such as power grid maintenance scheduling and professional car racing.

In the setting studied in this thesis, a learning problem provides estimated probabilities, and those probabilities all feed into a single optimization problem for decision-making. This is how people typically set up facility location problems, traveling repairman and traveling salesman problems, knapsack problems, and so on. The interesting part is how the uncertainty in predictions translates into uncertainty in the decision problem. We used tools from statistical learning theory and robust optimization to provide important insights into this question.

Machine learning with operational costs: Quantifying influence using statistical learning theory (Tulabandhula & Rudin, 2013, 2014b) - We proposed a way to align statistical modeling with decision making, which we called the "Machine Learning with Operational Costs (MLOC)" framework. We focused on the class of problems where we, (a) construct a data dependent predictive model (a classification or regression function), and (b) solve a decision making optimization problem whose parameters depend on the predictive model. We proposed a new method that propagates the uncertainty in predictive modeling to the uncertainty in operational cost, where operational cost is the amount spent by the practitioner in solving the problem. The method allows us to explore the range of operational costs associated with the set of reasonable statistical models as shown in Figure 1, so as to provide a useful way for practitioners to understand uncertainty. To do this computationally, we cast the operational cost as a regularization term in a learning algorithm's objective function, allowing either an optimistic or pessimistic view of possible costs, depending on the regularization parameter.

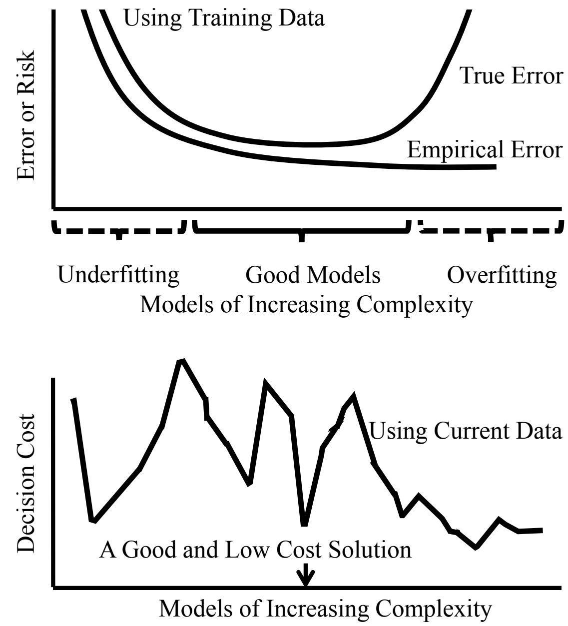

Figure 1: Operational/Decision costs change depending on the predictive model chosen.

From another perspective, if we have prior knowledge about the operational cost, for instance that it should be low, this knowledge can help to restrict the hypothesis space, and can help with generalization. In fact, it may be much more natural for a manager to have a prior belief on the cost to solve a problem than, for instance, a belief on the ℓ_2 norm of the coefficients of a predictive model, or the number of nonzero coefficients. Based on this view, we developed several statistical learning theory generalization bounds that take the operational cost knowledge into account. These bounds are novel, and require bounds on the complexity of a set of functions (covering numbers, VC dimension, Rademacher complexity, etc.). We also showed that learning with operational costs is related to the robust optimization framework.

Copyright © 2017 by the author(s).

We explored power grid maintenance using recent data provided by Con Edison for New York City (Tulabandhula & Rudin, 2014b), call center staffing, portfolio optimization and healthcare logistics problems under this framework. Our major findings were that it is very important to understand how the uncertainty in the predictive modeling translates into real-world uncertainty. For instance, for routing utility trucks on the NYC power grid, there can be quite a lot of uncertainty in what route is optimal for inspecting power grid equipment. This is something extremely important for power grid operators to know about; yet, they would not have been able to consider this uncertainty without our methods. Along similar lines and with similar challenges, we also performed knowledge discovery and designed a novel decision support tool for professional car racing (Tulabandhula & Rudin, 2014d).

Decision making backed by machine learning: Learning uncertainty sets for robust optimization (Tulabandhula & Rudin, 2014c) - We want our decision to best handle the the worst possible situation that could arise, out of an *uncertainty set* of possible situations. Classically, the uncertainty set is simply chosen by the user, or it might be estimated in overly simplistic ways with strong distributional assumptions; whereas in this work, we show how to *learn the uncertainty set from data collected in the past*. This method is principled, and backed by statistical learning theoretic bounds. And it will allow practitioners to construct uncertainty sets that are just right – not too conservative, and not too small.

Machine learning beyond decision making priors: An analysis using convex duality (Tulabandhula & Rudin, 2014a) - We considered a supervised learning setting where side knowledge is provided about the labels of an additional set of unlabeled examples. One of the ways such side knowledge can arise is through knowledge about an associated decision making problem (the MLOC setting). We considered many other sources of side knowledge than what had been studied rigorously in the past that lead to linear, polygonal, quadratic or conic constraints constraints on the hypothesis space. We proved bounds on the complexity measures of these constrained hypothesis spaces. These are some of the first results that enable a practitioner to understand the impact of different types of side knowledge on their learning problem.

Acknowledgments. Funding for this thesis came from a Fulbright Science and Technology Fellowship, the Solomon Buchsbaum Research Fund, and NSF grant IIS-1053407.

References

Tulabandhula, T., & Rudin, C. (2013). Machine Learning with Operational Costs. *Journal of Machine Learning Research*, *14*, 1989-2028.

Tulabandhula, T., & Rudin, C. (2014a). Generalization Bounds for Learning with Linear, Polygonal, Quadratic and Conic Side Knowledge. *Machine Learning Journal*, 1-34.

Tulabandhula, T., & Rudin, C. (2014b). On Combining Machine Learning with Decision Making. *Machine Learning Journal*, *97*(1-2), 33-64.

Tulabandhula, T., & Rudin, C. (2014c). Robust Optimization using Machine Learning for Uncertainty Sets. In *Proceedings of the International Symposium on Artificial Intelligence and Mathematics.*

Tulabandhula, T., & Rudin, C. (2014d). Tire Changes, Fresh Air, and Yellow Flags: Challenges in Predictive Analytics for Professional Racing. *Big Data*, *2*(2).

Theja Tulabandhula received his bachelors and masters degrees in electrical engineering from Indian Institute of Technology, Kharagpur, in 2009, where he was awarded the Prime Minister of India Gold medal. He received his Ph.D. degree in 2014 at the Massachusetts Institute of Technology (MIT), Cambridge. There, he was a Xerox-MIT fellow and a Science and Technology Fulbright scholar. He is currently a research scientist in the Analytics Lab at the Xerox Research Centre India.