Improving tuberculosis treatment by integrating optimization and learning

Bryan Wilder [1] **Jackson Killian** [1] **Amit Sharma** [2] **Vinod Choudhary** [3] **Bistra Dilkina** [4] **Milind Tambe** [1]

Abstract

Tuberculosis (TB) is one of the top ten causes of death worldwide, yet in most cases it is a curable and preventable disease. The prevalence of TB is caused in part by non-adherence to medication, which results in greater risk of death, reinfection and contraction of multidrug-resistant TB. Using data on 17,000 Indian patients provided by the NGO Everwell, we consider the problem of predicting which patients are likely to miss doses in the near future and optimizing interventions by health workers to avert such treatment failures. On a technical level, we propose a means of integrating common classes of discrete optimization problems into the training of deep learning or other predictive models. In the tuberculosis domain, we find that such decision-focused learning improves the number of successful interventions by approximately 15% compared to standard machine learning approaches, demonstrating that aligning the goals of learning and decision making can yield substantial benefits in a socially critical application.

1. Introduction

Tuberculosis (TB) is one of the top ten causes of death worldwide (WHO, 2018), despite widespread and effective treatment. One major challenge is non-adherence to medication, which results in greater risk of death, reinfection and contraction of multidrug-resistant TB (Thomas et al., 2005). Increasingly, health workers use *digital adherence technologies* (DATs) to determine when patients have taken their medicine (Subbaraman et al., 2018). We focus on improving adherence to tuberculosis treatment by leveraging digital adherence data, introducing new techniques at the intersection of optimization and learning. DATs allow patients to be "observed" consuming their medication electronically, e.g. via two-way text messaging, video capture, electronic pillboxes, or toll-free phone calls. Data from such devices enables health workers to triage patients and focus their limited resources on the highest risk patients. Preliminary studies suggest that DATs can improve adherence in multiple disease settings (Haberer et al., 2017; Corden et al., 2016), prompting its use and evaluation for managing TB adherence (Garfein et al., 2015; Liu et al., 2015; 99DOTS).

We study how the wealth of longitudinal data produced by DATs can be used to help health workers better triage TB patients and deliver interventions to boost overall adherence of their patient cohort. The data we analyze comes from a partnership with the nonprofit 99DOTS and the healthcare technology company Everwell who have implemented a DAT by which patients prove adherence through daily toll-free calls. 99DOTS operates in India where there were an estimated 2.7 million cases of TB in 2017 (WHO, 2018); they shared data from one major city in Maharashtra (referred to as "The City.") 99DOTS enables health workers to intervene with at-risk patients via texts, calls, or home visits. Note that many of these patients live in low-resource communities where each health worker manages tens to hundreds of patients; far more than they can possibly visit in a day. Thus, models that can identify patients at risk of missing doses and prioritize interventions by health workers are of paramount importance. We first propose the following prediction task: given adherence data up to a certain time period for patients not currently considered for intervention, predict risk of non-adherence in the next week and develop machine learning models. We then study a particular intervention task which requires workers to balance travel costs while predicting which patients will benefit most from interventions.

This task, like many other real-world uses of AI, requires a pipeline from data to decisions which involves two components: machine learning models and optimization algorithms. Our concern here is discrete optimization, which is ubiquitous in real-world applications of artificial intelligence (such as the TB domain). Typically, optimization and learning are treated entirely separately during training. That is, a system designer will first train a predictive model using some standard measure of accuracy Then, the model's predictions are given as input to the optimization algorithm

[1]School of Engineering and Applied Sciences, Harvard University [2]Microsoft Research India [3]RNTCP, Government of India [4]Department of Computer Science, University of Southern California. Correspondence to: Bryan Wilder <bryan.wilder0@gmail.com>.

Appearing at the International Conference on Machine Learning AI for Social Good Workshop, Long Beach, United States, 2019.

to produce a decision. Such *two-stage* approaches are extremely common across many domains. This process is justified when the predictive model is perfect, or near-so, since completely accurate predictions also produce the best decisions. However, in complex learning tasks, all models will make errors and the training process implicitly trades off where these errors will occur. When prediction and optimization are separate, this tradeoff is divorced from the goal of the broader pipeline: to make the best decision possible. We introduce methods which insert discrete optimization into the loop of machine learning training, enabling end-to-end training focused on the goal of improving interventions. In the tuberculosis domain, such decision-focused learning improves by about 15% over standard machine learning approaches, demonstrating the value of tailoring the learned model to fit the decision problem at hand.

2. Previous work

Adherence tracking and prediction: Outcomes and adherence research are well studied in the medical literature (Kardas et al., 2013). Traditionally, studies have attempted to identify demographic or behavioral factors correlated with non-adherence so that health workers can focus interventions on patients who are likely to fail. Typically these studies gather demographic and medical statistics on a cohort of patients, observe the cohort's adherence and outcomes throughout the trial, then retrospectively apply survival (Shargie & Lindtjørn, 2007; Kliiman & Altraja, 2010) or logistic regression (Roy et al., 2015) analysis to determine covariates predictive of failure. Newer work has improved classification accuracy via machine learning techniques (Kalhori & Zeng, 2013; Hussain & Junejo, 2018; Sauer et al., 2018; Mburu et al., 2018). However, the conclusions connecting predictors to risk are largely the same as in previous medical literature. While such studies have improved patient screening at the time of diagnosis, they offer little knowledge about how risk changes *during* treatment. In this work, we show how a patient's real-time adherence data can be used to track and predict risk changes throughout the course of their treatment.

Machine learning: There is a growing body of research at the interface of machine learning and discrete optimization (Vinyals et al., 2015; Bertsimas & Dunn, 2017; Khalil et al., 2017b;a). However, previous work largely focuses on either using discrete optimization to find an accuracy-maximizing predictive model or using machine learning to speed up optimization algorithms. Here, we pursue a deeper synthesis; to our knowledge, this work is the first to train predictive models using combinatorial optimization performance with the goal of improving decision making. The closest work to ours in motivation is (Donti et al., 2017), who study task-based convex optimization. Their aim is to optimize a convex function which depends on a learned parameter. As in their work, we use the idea of differentiating through the KKT conditions. However, their focus is entirely on continuous problems. Our discrete setting raises new technical challenges, highlighted below.

3. Data Description

99DOTS provides each patient with a cover for their sleeve of pills that associates a hidden unique phone number with each pill. As patients expose each pill, they expose the associated phone number. Each patient is instructed to place a toll-free call to the indicated number each day. 99DOTS counts a dose only if the patient calls the correct number for a given day. Due to the sensitivity of the health domain, all data provided by our partners was fully anonymized before we received it. The dataset contains over 2.1 million calls by about 17,000 patients, served by 252 health centers across The City.

Patient Details. This is the primary record for patients who have enrolled with 99DOTS. The table includes demographic features such as weight-band, age-band, gender and treatment center ID. Also included are treatment start and end dates, whether treatment is completed or ongoing, and an "adherence string" which summarizes a patient's daily adherence. For patients who completed treatment, a treatment outcome is also assigned according to the standard WHO definitions (WHO, 2013).

Mapping phone numbers to patients. Patients must call from a registered phone number for a dose to be counted by the 99DOTS system. Patients can register multiple phones, each of which will be noted in the Phone Map table. We filtered out phones that were registered to multiple patients since they could not be uniquely mapped to patients. Also, patients who had *any* calls from shared phones were filtered out to avoid analyzing incomplete call records. This removed $< 1\%$ of the patients from the data set.

4. Problem description

We focus on a specific optimization problem that models the allocation of health workers to intervene with patients who are at risk over the course of a week. The health worker is responsible for a population of patients across different locations, and may visit one location each day. We use location identifiers at the level of the TB Unit since this is the most granular identifier which is shared by the majority of patients in our dataset. Visiting a location allows the health worker to intervene with any of the patients at that location. The optimization problem is to select a set of locations to visit which maximizes the number of patients who receive an intervention *on or before the first day they would have missed a dose*. We refer to this quantity as the number of *successful interventions*; this captures the extent to which health workers are able to intervene before problems arise.

We now show how this optimization problem can be formalized as a linear program. We have a set of locations $i = 1...L$ and patients $j = 1...N$ where patient j has location ℓ_j. Over days of the week $t = 1...7$, the objective coefficient p_{jt} is 1 if an intervention on day t with patient j is successful and 0 otherwise. Our decision variable is x_{it}, and takes the value 1 if the health worker visit location i on day t and 0 otherwise. With this notation, the final LP is as follows:

$$
\begin{aligned}
\max_x \; & \sum_{t=1}^{7} \sum_{i=1}^{L} x_{it} \left(\sum_{j:\ell_j = i} p_{jt} \right) \\
\text{s.t.} \; & \sum_{i=1}^{L} x_{it} \leq 1, t = 1...7 \\
& \sum_{t=1}^{7} x_{it} \leq 1, i = 1...L
\end{aligned}
$$

where the second constraint prevents the objective from double-counting multiple visit to a location. We remark that the feasible region of the LP can be shown to be equivalent to a bipartite matching polytope, implying that the optimal solution is always integral.

5. Machine learning approach

The machine learning task is to predict the values of the p_{jt}, which are unknown at the start of the week. Two sources of data are available for each patient. First, the patient's adherence data over the previous work (as provided by the 99DOTS system). Second, a set of demographic features such as weight-band, age-band, gender and treatment center ID. We develop a combined neural network architecture which uses an LSTM to process the adherence time series from the previous week, and combines the hidden state of the LSTM with the demographic features through fully connected layer. We refer to the final model as DeepNet. We develop two variants of DeepNet. First, one trained using a standard approach: cross-entropy loss. Second, one which incorporates the above LP into the training process. Specifically, our goal is to solve the LP in each forward pass of the training process and then differentiate through the solution to the LP (as a function of the p_{jt}) in the backward pass. This allows us to use the objective value of the LP solution *with respect to the ground truth* p_{jt} as the loss function for training. This *decision-focused* training method incentivizes the machine learning model to focus specifically what is needed to make good decisions.

More formally, let $x(p)$ denote the optimal solution the LP with respect to the predictions p. In order to train end-to-end, we would like to compute the derivative $\frac{\partial x}{\partial p}$ so that we can pass gradients through the optimization problem. The optimal decision $x(p)$ must satisfy the KKT conditions, which define a system of linear equations based on the gradients of the objective and constraints around the optimal point. Is is known that by applying the implicit function theorem, we can differentiate the solution to this linear system (Gould et al., 2016; Donti et al., 2017). In more detail, suppose that we have an LP in standard form, maximizing $c^T x$ subject to the constraint $x \in \mathcal{X} = \{Ax \leq b\}$ for some matrix A and vector b. Let (x, λ) be pair of primal and dual variables which satisfy the KKT conditions. Then differentiating the conditions yields that

$$
\begin{bmatrix} \nabla_x^2 c^T x & A^T \\ diag(\lambda)A & diag(Ax - b) \end{bmatrix} \begin{bmatrix} \frac{dx}{dc} \\ \frac{d\lambda}{dc} \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix} \qquad (1)
$$

By solving this system of linear equations, we can obtain the desired term $\frac{dx}{dc}$. This approach runs into an immediate difficulty: the optimal solution to an LP may not be differentiable (or even continuous) with respect to c. This is because the optimal solution may "jump" to a different vertex. Formally, the left-hand side matrix in Equation 1 becomes singular since the Hessian $\nabla_x^2 c^T x$ is always zero. We resolve this challenge by instead solving the regularized problem

$$
\max c^T x - \gamma ||x||_2^2 \;\text{ s.t. }\; Ax = b \qquad (2)
$$

which introduces a penalty proportional to the squared norm of the decision vector. This transforms the LP into a strongly concave quadratic program (QP). The Hessian is just $-2\gamma I$ (where I is the identity matrix), which renders the solution differentiable under mild conditions (see the supplement for proofs[1]):

Theorem 1. *Provided that the LP is feasible and all rows of A are linearly independent, $x(c)$ is differentiable with respect to c almost everywhere. If A has linearly dependent rows, removing these rows yields an equivalent problem which is differentiable almost everywhere. Wherever $x(c)$ is differentiable, it satisfies the conditions in Equation 1.*

Moreover, we can control the loss that regularization can cause on the original, linear problem:

[1] https://www.dropbox.com/s/v3tc6tk6qn6byyz/supplement_tb_icml.pdf?dl=0

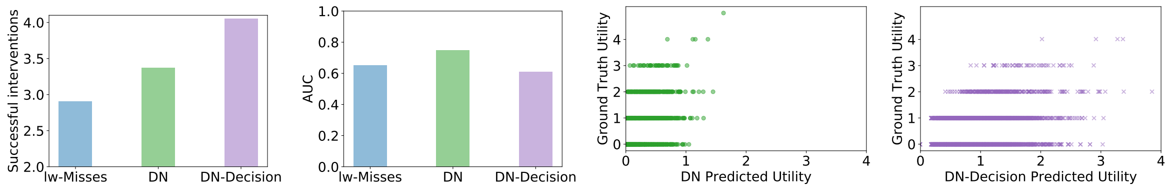

Figure 1. Experimental results. From left to right: average missed doses averted by each method, AUC of each method, scatter plot of predicted utility vs ground truth for the DN model, and the same scatter plot for DN-Decision.

Theorem 2. *Define* $D = \max_{x,y\in\mathcal{X}} ||x-y||^2$ *and* OPT *to be the optimal value for the LP. We have* $c^\top x(c) \geq OPT - \gamma D$.

Together, these results give us a differentiable surrogate which still enjoys an approximation guarantee relative to the integral problem. At test time, we simply set $\gamma = 0$ to produce an integral decision. We remark that while this paper focuses on the specific motivating domain of TB, our techniques apply more generally to other combinatorial LPs.

6. Experimental results

We compare three models. First, a baseline model which approximates the strategy that health workers use to prioritize patients in the status quo (essentially, intervening with those who have recently missed more than some number of doses). Specifically, we threshold the number of doses patient j missed in the last week, setting $c_{jt} = 0$ for all t if this value falls below the threshold τ and $c_{jt} = 1$ otherwise. We used $\tau = 1$ since it performed best. This strategy is refered to last week misses (lw-Misses). Second, we trained our DeepNet system (DN) directly on the true c_{jt} as a binary prediction task using cross-entropy loss. Third, we trained DeepNet to predict c_{jt} using decision-focused learning. We refer to this model as DN-Decision. We created instances of the decision problem by randomly partitioning patients into groups of 100, modeling a health worker under severe resource constraints.

Figure 1 shows results for this task. In the top row, we see that DN and DN-Decision both outperform lw-Misses, as expected. DN-Decision improves the number of successful interventions by approximately 15% compared to DN, demonstrating the value of tailoring the learned model to a given planning problem. DN-Decision actually has worse AUC than either DN or lw-Misses, indicating that typical measures of machine learning accuracy are not a perfect proxy for utility in decision making. To investigate what specifically distinguishes the predictions made by DN-Decision, the bottom row of Figure 1 shows scatter plots of the predicted utility at each location according to DN and DN-Decision versus the true values. Visually, DN-Decision appears better able to distinguish the high-utility outliers which are most important to making good decisions. Quantitatively, DN-Decision's predictions have worse correlation with the ground truth overall (0.463, versus 0.519 for DN), but better correlation on locations where the true utility is strictly more than 1 (0.504 versus 0.409). Hence, decision-focused training incentivizes the model to focus on making accurate predictions specifically for locations that are likely to be good candidates for an intervention. This demonstrates the benefit of our flexible machine learning modeling approach, which can use custom-defined loss functions to automatically adapt to particular decision problems.

7. Conclusion

We present a framework for learning to make intervention recommendations from data generated by digitial adherence systems applied to TB care. Using patient data from India provided by the nonprofit 99DOTS, we train sequence prediction models to forecast patient adherence and allow health workers to better target their interventions. Further, we develop techniques to specialize the training of such models to particular decision problems (in particular, discrete optimization problems represented as linear programs). These techniques are applied in the context of a specific decision problem that models the tradeoff between travel time and patient risk for a resource-constrained health worker. We show that tailoring our model for a specific intervention via decision-focused learning can improve performance by a further 15%. The learning approaches we present here are general and could be leveraged to study data generated by DATs as applied to any medication regimen. With the growing popularity of DAT systems for TB, HIV, Diabetes, Heart Disease, and other medications, we hope to lay the groundwork for improved patient outcomes in healthcare settings around the world.

References

99DOTS. 99dots. URL https://www.99dots.org/.

Bertsimas, D. and Dunn, J. Optimal classification trees. *Machine Learning*, 106(7):1039–1082, 2017.

Corden, M. E., Koucky, E. M., Brenner, C., Palac, H. L., Soren, A., Begale, M., Ruo, B., Kaiser, S. M., Duffecy, J., and Mohr, D. C. Medlink: A mobile intervention to improve medication adherence and processes of care for treatment of depression in general medicine. *Digital Health*, 2:2055207616663069, 2016.

Donti, P., Amos, B., and Kolter, J. Z. Task-based end-to-end model learning in stochastic optimization. In *Advances in Neural Information Processing Systems*, pp. 5484–5494, 2017.

Garfein, R. S., Collins, K., Muñoz, F., Moser, K., Cerecer-Callu, P., Raab, F., Rios, P., Flick, A., Zúñiga, M. L., Cuevas-Mota, J., et al. Feasibility of tuberculosis treatment monitoring by video directly observed therapy: a binational pilot study. *The International Journal of Tuberculosis and Lung Disease*, 19(9):1057–1064, 2015.

Gould, S., Fernando, B., Cherian, A., Anderson, P., Cruz, R. S., and Guo, E. On differentiating parameterized argmin and argmax problems with application to bi-level optimization. *arXiv preprint arXiv:1607.05447*, 2016.

Haberer, J. E., Musinguzi, N., Tsai, A. C., Bwana, B., Muzoora, C., Hunt, P., Martin, J., Bangsberg, D., et al. Real-time electronic adherence monitoring plus follow-up improves adherence compared with standard electronic adherence monitoring. *AIDS (London, England)*, 31(1): 169–171, 2017.

Hussain, O. A. and Junejo, K. N. Predicting treatment outcome of drug-susceptible tuberculosis patients using machine-learning models. *Informatics for Health and Social Care*, pp. 1–17, 2018.

Kalhori, S. R. N. and Zeng, X.-J. Evaluation and comparison of different machine learning methods to predict outcome of tuberculosis treatment course. *Journal of Intelligent Learning Systems and Applications*, 5(03):184, 2013.

Kardas, P., Lewek, P., and Matyjaszczyk, M. Determinants of patient adherence: a review of systematic reviews. *Frontiers in pharmacology*, 4:91, 2013.

Khalil, E. B., Dai, H., Zhang, Y., Dilkina, B., and Song, L. Learning combinatorial optimization algorithms over graphs. In *NIPS*, 2017a.

Khalil, E. B., Dilkina, B., Nemhauser, G. L., Ahmed, S., and Shao, Y. Learning to run heuristics in tree search. In *IJCAI*, 2017b.

Kliiman, K. and Altraja, A. Predictors and mortality associated with treatment default in pulmonary tuberculosis. *The international journal of tuberculosis and lung disease*, 14(4):454–463, 2010.

Liu, X., Lewis, J. J., Zhang, H., Lu, W., Zhang, S., Zheng, G., Bai, L., Li, J., Li, X., Chen, H., et al. Effectiveness of electronic reminders to improve medication adherence in tuberculosis patients: a cluster-randomised trial. *PLoS medicine*, 12(9):e1001876, 2015.

Mburu, J. W., Kingwara, L., Ester, M., and Andrew, N. Use of classification and regression tree (cart), to identify hemoglobin a1c (hba1c) cut-off thresholds predictive of poor tuberculosis treatment outcomes and associated risk factors. *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*, 11:10–16, 2018.

Roy, N., Basu, M., Das, S., Mandal, A., Dutt, D., and Dasgupta, S. Risk factors associated with default among tuberculosis patients in darjeeling district of west bengal, india. *Journal of family medicine and primary care*, 4(3): 388, 2015.

Sauer, C. M., Sasson, D., Paik, K. E., McCague, N., Celi, L. A., Fernandez, I. S., and Illigens, B. M. Feature selection and prediction of treatment failure in tuberculosis. *PloS one*, 13(11):e0207491, 2018.

Shargie, E. B. and Lindtjørn, B. Determinants of treatment adherence among smear-positive pulmonary tuberculosis patients in southern ethiopia. *PLoS medicine*, 4(2):e37, 2007.

Subbaraman, R., de Mondesert, L., Musiimenta, A., Pai, M., Mayer, K. H., Thomas, B. E., and Haberer, J. Digital adherence technologies for the management of tuberculosis therapy: mapping the landscape and research priorities. *BMJ global health*, 3(5):e001018, 2018.

Thomas, A., Gopi, P., Santha, T., Chandrasekaran, V., Subramani, R., Selvakumar, N., Eusuff, S., Sadacharam, K., and Narayanan, P. Predictors of relapse among pulmonary tuberculosis patients treated in a dots programme in south india. *The International Journal of Tuberculosis and Lung Disease*, 9(5):556–561, 2005.

Vinyals, O., Fortunato, M., and Jaitly, N. Pointer networks. In *NIPS*, 2015.

WHO. *Definitions and reporting framework for tuberculosis 2013 revision*. World Health Organization, Geneva, 2013. ISBN 978 92 4 150534 5.

WHO. *Global tuberculosis report 2018*. World Health Organization, 2018. Licence: CC BY-NC-SA 3.0 IGO.