Integrating optimization and learning to prescribe interventions for tuberculosis patients

Bryan Wilder[1], Jackson A. Killian[1], Amit Sharma[2], Vinod Choudhary[3], Bistra Dilkina[1], Milind Tambe[1]

[1]University of Southern California, {bwilder, jakillia, dilkina, tambe}@usc.edu
[2] Microsoft Research India, amshar@microsoft.com
[3]RNTCP, Govt. of India, drvinod06@gmail.com

Abstract. Creating impact in real-world settings requires agents which navigate the full pipeline from data, to predictive models, to decisions. These components are typically approached separately: a machine learning model is first trained via a measure of predictive accuracy, and then its predictions are used as input into an optimization algorithm which produces a decision. However, the loss function used to train the model may easily be misaligned with the end goal of the agent, which is to make the best decisions possible.

We focus on combinatorial optimization problems and introduce a general framework for decision-focused learning, where the machine learning model is directly trained in conjunction with the optimization algorithm to produce high-quality decisions. Technically, our contribution is a means of integrating common classes of discrete optimization problems into deep learning or other predictive models, which are typically trained via gradient descent. The main idea is to use a continuous relaxation of the discrete problem to propagate gradients through the optimization procedure. We instantiate this framework for two broad classes of combinatorial problems: linear programs and submodular maximization.

We then provide an application of such techniques to a real problem of societal importance: improving interventions in tuberculosis treatment. 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 of medication in the near future and optimizing interventions by health workers to avert such treatment failures. We find the 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]

Keywords: Machine learning · Optimization · Tuberculosis

[1] This paper combines material from two sources. First, [48], which presents the technical approach to decision-focused combinatorial optimization. Second, [28], which develops the machine learning approach to tuberculosis adherence prediction and applies the techniques of [48]. The present work combines the two as a complete case study of integrating learning and optimization for a real-world societal challenge.

1 Introduction

The goal in many real-world applications of artificial intelligence is to create a pipeline from data, to predictive models, to decisions. Together, these steps enable a form of evidence-based decision making which has transformative potential across domains such as healthcare, scientific discovery, transportation, and more [20, 19]. This pipeline requires two technical components: machine learning models and optimization algorithms, which must function together as part of a combined agent. Machine learning models use the data to predict unknown quantities; optimization algorithms use these predictions to arrive at a decision which maximizes some objective. Our concern here is combinatorial optimization, which is ubiquitous in real-world applications of artificial intelligence, ranging from matching applicants to public housing to selecting a subset of movies to recommend. We focus on common classes of combinatorial problems which have well-structured continuous relaxations, e.g., linear programs and submodular maximization. A vast literature has been devoted to combinatorial optimization [31]. Importantly though, optimization is often insufficient without the broader pipeline because the objective function is unknown and must predicted via machine learning.

While machine learning has witnessed incredible growth in recent years, the two pieces of the pipeline are treated entirely separately by typical training approaches. That is, a system designer will first train a predictive model using some standard measure of accuracy, e.g., mean squared error for a regression problem. Then, the model's predictions are given as input to the optimization algorithm to produce a decision. Such *two-stage* approaches are extremely common across many domains [47, 13, 33, 49]. 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 agent: to make the best decision possible.

We propose a *decision-focused learning* framework which integrates the machine learning and optimization components of an agent by training both as a single end-to-end system. That is, the predictive model is trained using the quality of the decisions which it induces via the optimization algorithm. Similar ideas have recently been explored in the context of convex optimization [11], but to our knowledge ours is the first attempt to train machine learning systems for performance on *combinatorial* decision-making problems. Combinatorial settings raise new technical challenges because the optimization problem is discrete. However, machine learning systems (e.g., deep neural networks) are often trained via gradient descent.

Our first contribution is a general framework for training machine learning models via their performance on combinatorial problems. The starting point is to relax the combinatorial problem to a continuous one. Then, we analytically differentiate the optimal solution to the continuous problem as a function of the model's predictions. This allows us to train using a continuous proxy for the discrete problem. At test time, we round the continuous solution to a discrete point.

Our second contribution is to instantiate this framework for two broad classes of combinatorial problems: linear programs and submodular maximization problems. Linear programming encapsulates a number of classical problems such as shortest path, maximum flow, and bipartite matching. Submodular maximization, which reflects the intuitive phenomena of diminishing returns, is also ubiquitous; applications range from social networks [25] to recommendation systems [45]. In each case, we resolve a set of technical challenges to produce well-structured relaxations which can be efficiently differentiated through.

Our third contribution is to present a case study of how such techniques can be applied towards a real societal challenge. We focus on improving adherence to tuberculosis treatment by leveraging digital adherence data and decision-focused learning. The World Health Organization (WHO) reports that the lung disease tuberculosis (TB) is one of the top ten causes of death worldwide [35], 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 [42]. To combat non-adherence, the WHO standard protocol is Directly Observed Treatment, short-course (DOTS), in which a health worker directly observes and confirms that a patient is consuming the required medication multiple times in a week. However, requiring patients to travel to the DOTS clinic causes financial burden, and potentially social stigma due to public fear of the disease. Such barriers cause patients to default from treatment, making TB eradication difficult. Thus, digital adherence technologies (DATs), which give patients flexible means to prove adherence, have gained popularity globally [41].

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. Health workers can then view real-time patient adherence on a dashboard. In addition to improving patient flexibility and privacy, the dashboard 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 [17, **?**,**?**], prompting its use and evaluation for managing TB adherence [15, 32, **?**]. The WHO has even published a guide for the proper implementation of the technology in TB care [36].

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 [1] and the healthcare technology company Everwell [12] 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 [35]; they shared data from one major city in Maharashtra (referred to as "The City.") Patients enrolled in 99DOTS in The City currently receive interventions according to the following general guidelines. If they have not taken their medication by the afternoon, they (and their health worker) receive a text message reminder. If the patient still does not take their medication by some time later, the worker will call the patient directly. Finally, if a patient simply does not respond to these previous interventions after some number of days, they may be personally visited by a health worker. 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. *In this setting, 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. With our proposed models, 99DOTS can now leverage several years of collected adherence data to better inform patient care and prioritize limited intervention resources.

2 Previous work

2.1 Machine learning and optimization

There is a growing body of research at the interface of machine learning and discrete optimization [46, 4, 27, 26]. 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 [11], 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. Elmachtoub and Grigas elmachtoub2017smart also propose a means of integrating prediction and optimization; however, their method applies strictly to linear optimization and focuses on linear predictive models while our framework applies to nonlinear problems with more general models (e.g., neural networks). Finally, some work has noted that two-stage methods lead to poor optimization performance in specific domains [5, 14].

Our work is also related to recent research in structured prediction [3, 44, 34, 9]. which aims to make a prediction lying in a discrete set. This is fundamentally different than our setting since their goal is to *predict* an external quantity, not to *optimize* and find the best decision possible. However, structured prediction sometimes integrates a discrete optimization problem as a module within a larger neural network. The closest such work technically to ours is [43], who design a differentiable algorithm for submodular maximization in order to predict choices made by users. Their approach is to introduce noise into the standard greedy algorithm, making the probability of outputting a given set differentiable. There are two key differences between our approaches. First, their approach does not apply to the decision-focused setting because it maximizes the likelihood of a *fixed* set but cannot optimize for finding the best set. Second, exactly computing gradients for their algorithm requires marginalizing over the $k!$ possible permutations of the items, forcing a heuristic approximation to the gradient. Our approach allows closed-form differentiation.

Some deep learning architectures differentiate through gradient descent steps, related to our approach in the submodular setting. Typically, previous approaches explicitly unroll T iterations of gradient descent in the computational graph [10]. However, this approach is usually employed for *unconstrained* problems where each iteration is a simple gradient step. By contrast, our combinatorial problems are constrained, requiring a projection step to enforce feasibility. Unrolling the projection step may be difficult, and would incur a large computational cost. We instead exploit the fact that gradient ascent converges to a local optimum and analytically differentiate via the KKT conditions.

2.2 Adherence tracking and prediction

Outcomes and adherence research are well studied in the medical literature for a variety of diseases [24]. 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. Tuberculosis in particular, given its lethality and prevalence in third world countries, has been studied throughout the world including in Ethiopia [40], Estonia [30], and India [39]. 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 [40, **?**] or logistic regression [39] analysis to determine covariates predictive of failure. Newer work has improved classification accuracy via machine learning techniques such as Decision Trees, Neural Networks, Support Vector Machines and more [23, **?**,**?**,**?**]. 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. Previous studies likely did not address this question because accurately measuring patient adherence has historically been difficult.

However, in recent years, new technologies have made measuring daily adherence feasible in the context of many diseases such as HIV or stroke. One such common device is an electronic pill bottle cap that records the date/time when the cap is removed. While some previous work has used electronic cap data to determine predictors of non-adherence [38, 37, 8], almost no research has used the daily measurements made possible by the electronic cap to study changes in adherence over time. One study used data from a smart pillbox to *retrospectively* categorize patient adherence [29], but our focus is on *prospective* identification of patients at risk of missing doses before failures occur. As such devices enter mainstream use, machine learning techniques like the ones that we propose will play an important role in the treatment of a wide spectrum of diseases.

3 Problem description

We start out by introducing our general technical approach to integrating learning an optimization; application to the tuberculosis domain will be considered in Section 5. We study combinatorial optimization problems of the form $\max_{x \in \mathcal{X}} f(x, \theta)$, where $\mathcal{X}$ is a discrete set enumerating the feasible decisions. Without loss of generality, $\mathcal{X} \subseteq \{0, 1\}^n$ and the decision variable x is a binary vector. The objective f depends on a parameter $\theta \in \Theta$. If θ were known exactly, a wide range of existing techniques could be used to solve the problem. In this paper, we consider the challenging (but prevalent) case where θ is unknown and must be inferred from data. For instance, in bipartite matching, x represents whether each pair of nodes were matched and θ contains the reward for matching each pair. In many applications, these affinities are learned from historical data.

Specifically, the decision maker observes a feature vector $y \in \mathcal{Y}$ which is correlated with θ. This introduces a learning problem which must be solved prior to optimization. As in classical supervised learning, we formally model y and θ as drawn from a joint distribution P. Our algorithm will observe training instances $(y_1, \theta_1)...(y_N, \theta_N)$ drawn iid from P. At test time, we are give a feature vector y corresponding to an *unobserved* θ. Our algorithm will use y to predict a parameter value $\hat{\theta}$. Then, we will solve the optimization problem $\max_x f(x, \hat{\theta})$ to obtain a decision x^*. Our utility is the objective value that x^* obtains with respect to the *true but unknown* parameter θ, $f(x^*, \theta)$.

Let $m : \mathcal{Y} \to \Theta$ denote a model mapping observed features to parameters. Our goal is to (using the training data) find a model m which maximizes expected performance on the underlying optimization task. Define $x^*(\theta) = \arg\max_{x \in \mathcal{X}} f(x, \theta)$ to be the optimal x for a given θ. The end goal of the data-decisions pipeline is to maximize

$$\mathop{\mathbb{E}}_{y, \theta \sim P} [f(x^*(m(y)), \theta)] \tag{1}$$

The classical approach to this problem is a *two-stage* method which first learns a model using a task-agnostic loss function (e.g., mean squared error) and then uses the learned model to solve the optimization problem. The model class will have its own parameterization, which we denote by $m(y, \omega)$. For instance, the model class could consist of deep neural networks where ω denotes the weights. The two-stage approach first solves the problem $\min_\omega \mathbb{E}_{y,\theta\sim P}\left[\mathcal{L}(\theta, m(y,\omega))\right]$, where $\mathcal{L}$ is a loss function. Such a loss function measures the overall "accuracy" of the model's predictions but does not specifically consider how m will fare when used for decision making. The question we address is whether it is possible to do better by specifically training the model to perform well on the decision problem.

4 General framework

Our goal is to integrate combinatorial optimization into the loop of gradient-based training. That is, we aim to directly train the predictive model m by running gradient steps on the objective in Equation 1, which integrates both prediction and optimization. The immediate difficulty is the dependence on $x^*(m(y,\omega))$. This term is problematic for two reasons. First, it is a discrete quantity since x^* is a decision from a binary set. This immediately renders the output nondifferentiable with respect to the model parameters ω. Second, even if x^* were continuous, it is still defined as the solution to an optimization problem, so calculating a gradient requires us to differentiate through the argmax operation.

We resolve both difficulties by considering a continuous relaxation of the combinatorial decision problem. We show that for a broad class of combinatorial problems, there are appropriate continuous relaxations such that we can analytically obtain derivatives of the continuous optimizer with respect to the model parameters. This allows us to train any differentiable predictive model via gradient descent on a continuous surrogate to Equation 1. At test time, we solve the true discrete problem by rounding the continuous point.

More specifically, we relax the discrete constraint $x \in \mathcal{X}$ to the continuous one $x \in conv(\mathcal{X})$ where $conv$ denotes the convex hull. Let $x(\theta) = \arg\max_{x\in conv(\mathcal{X})} f(x,\theta)$ denote the optimal solution to the continuous problem. To train our predictive model, we would like to compute gradients of the whole-pipeline objective given by Equation 1, replacing the discrete quantity x^* with the continuous x. We can obtain a stochastic gradient estimate by sampling a single (y,θ) from the training data. On this sample, the chain rule gives

$$\frac{df(x(\hat{\theta}),\theta)}{d\omega} = \frac{df(x(\hat{\theta}),\theta)}{dx(\hat{\theta})}\frac{dx(\hat{\theta})}{d\hat{\theta}}\frac{d\hat{\theta}}{d\omega}$$

The first term is just the gradient of the objective with respect to the decision variable x, and the last term is the gradient of the model's predictions with respect to its own internal parameterization.

The key is computing the middle term, which measures how the optimal decision changes with respect to the prediction $\hat{\theta}$. For continuous problems, the optimal continuous decision x must satisfy the KKT conditions (which are sufficient for convex problems). The KKT conditions 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 [16, 11]. In more detail, recall that our continuous problem is over $conv(\mathcal{X})$, the convex hull of the discrete feasible solutions. This set is a polytope, which can be represented via linear equalities as the set $\{x : Ax \leq b\}$ 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

$$\begin{bmatrix} \nabla_x^2 f(x,\theta) & A^T \\ diag(\lambda)A & diag(Ax-b) \end{bmatrix} \begin{bmatrix} \frac{dx}{d\theta} \\ \frac{d\lambda}{d\theta} \end{bmatrix} = \begin{bmatrix} \frac{d\nabla_x f(x,\theta)}{d\theta} \\ 0 \end{bmatrix} \tag{2}$$

By solving this system of linear equations, we can obtain the desired term $\frac{dx}{d\theta}$. However, the above approach is a general framework; our main technical contribution is to instantiate it for specific classes of combinatorial problems. Specifically, we need (1) an appropriate continuous relaxation, along with a means of solving the continuous optimization problem and (2) efficient access to the terms in Equation 2 which are needed for the backward pass (i.e., gradient computation). We provide both ingredients for two broad classes of problems: linear programming and submodular maximization. In each setting, the high-level challenge is to ensure that the continuous relaxation is differentiable, a feature not satisfied by naive alternatives. We also show how to efficiently compute terms needed for the backward pass, especially for the more intricate submodular case.

4.1 Linear programming

The first setting that we consider is combinatorial problems which can be expressed as a linear program with equality and inequality constraints in the form

$$\max \theta^T x \text{ s.t. } Ax = b,\ Gx \leq h \tag{3}$$

Example problems include shortest path, maximum flow, bipartite matching, and a range of other domains. For instance, in a shortest path problem θ contains the cost for traversing each edge, and we are interested in problems where the true costs are unknown and must be predicted. Since the LP can be regarded as a continuous problem (it just happens that the optimal solutions in these example domains are integral), we could attempt to apply Equation 2 and differentiate the solution. This approach runs into an immediate difficulty: the optimal solution to an LP may not be differentiable (or even continuous) with respect to θ. This is because the optimal solution may "jump" to a different vertex. Formally, the left-hand side matrix in Equation 2 becomes singular since $\nabla_x^2 f(x,\theta)$ is always zero. We resolve this challenge by instead solving the regularized problem

$$\max \theta^T x - \gamma||x||_2^2 \text{ s.t. } Ax = b,\ Gx \leq h \tag{4}$$

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 given by $\nabla_x^2 f(x, \theta) = -2\gamma I$ (where I is the identity matrix), which renders the solution differentiable under mild conditions (see supplement for proof):

Theorem 1. *Let $x(\theta)$ denote the optimal solution of Problem 4. Provided that the problem is feasible and all rows of A are linearly independent, $x(\theta)$ is differentiable with respect to θ almost everywhere. If A has linearly dependent rows, removing these rows yields an equivalent problem which is differentiable almost everywhere. Wherever $x(\theta)$ is differentiable, it satisfies the conditions in Equation 2.*

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

Theorem 2. *Define $D = \max_{x,y \in conv(\mathcal{X})} ||x-y||^2$ as the squared diameter of the feasible set and OPT to be the optimal value for Problem 3. We have $\theta^\top x(\theta) \geq OPT - \gamma D$.*

Together, these results give us a differentiable surrogate which still enjoys an approximation guarantee relative to the integral problem. Computing the backward pass via Equation 2 is now straightforward since all the relevant terms are easily available. Since $\nabla_x \theta^\top x = \theta$, we have $\frac{d\nabla_x f(x,\theta)}{d\theta} = I$. All other terms are easily computed from the optimal primal-dual pair (x, λ) which is output by standard QP solvers. We can also leverage a recent QP solver [2] which maintains a factorization of the KKT matrix for a faster backward pass. At test time, we simply set $\gamma = 0$ to produce an integral decision.

4.2 Submodular maximization

We consider problems where the underlying objective to maximize a set function $f : 2^V \rightarrow R$, where V is a ground set of items. A set function is *submodular* if for any $A \subseteq B$ and any $v \in V \setminus B$, $f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B)$. We will restrict our consideration to submodular functions which are *monotone* $(f(A \cup \{v\}) - f(A) \geq 0 \; \forall A, v)$ and *normalized* $f(\emptyset) = 0$. This class of functions contains many combinatorial problems which have been considered in machine learning and artificial intelligence (e.g., influence maximization, facility location, diverse subset selection, etc.). We focus on the cardinality-constrained optimization problem $\max_{|S| \leq k} f(S)$, though our framework easily accommodates more general matroid constraints.

Continuous relaxation: We employ the canonical continuous relaxation for submodular set functions, which associates each set function f with its *multilinear extension* F [7]. We can view a set function as defined on the domain $\{0, 1\}^{|V|}$, where each element is an indicator vector which the items contained in the set. The extension F is a continuous function defined on the hypercube $[0, 1]^{|V|}$. We interpret a given fraction vector $x \in [0, 1]^{|V|}$ as giving the marginal probability that each item is included in the set. $F(x)$ is the expected value of $f(S)$ when each item i is included in S independently with probability x_i. In other words, $F(x) = \sum_{S \subseteq V} f(S) \prod_{i \in S} x_i \prod_{i \not\in S} 1 - x_i$. While this definition sums over exponentially many terms, arbitrarily close approximations can be obtained via random sampling. Further, closed forms are available for many cases of interest [21]. Importantly, well-known rounding algorithms [7] can convert a fractional point x to a set S satisfying $\mathbb{E}[f(S)] \geq F(x)$; i.e., the rounding is lossless.

As a proxy for the discrete problem $\max_{|S| \leq k} f(S)$, we can instead solve $\max_{x \in conv(\mathcal{X})} F(x)$, where $\mathcal{X} = \{x \in \{0,1\}^{|V|} : \sum_i x_i \leq k\}$. Unfortunately, F is not in general concave. Nevertheless, many first-order algorithms still obtain a constant factor approximation. For instance, a variant of the Frank-Wolfe algorithm solves the continuous maximization problem with the optimal approximation ratio of $(1 - 1/e)$ [7, 6].

However, non-concavity complicates the problem of differentiating through the continuous optimization problem. Any polynomial-time algorithm can only be guaranteed to output a *local* optimum, which need not be unique (compared to strongly convex problems, where there is a single global optimum). Consequently, the algorithm used to select $x(\theta)$ might return a *different* local optimum under an infinitesimal change to θ. For instance, the Frank-Wolfe algorithm (the most common algorithm for continuous submodular maximization) solves a linear optimization problem at each step. Since (as noted above), the solution to a linear problem may be discontinuous in θ, this could render the output of the optimization problem nondifferentiable.

We resolve this difficulty through a careful choice of optimization algorithm for the forward pass. Specifically, we use apply projected stochastic gradient ascent (SGA), which has recently been shown to obtain a $\frac{1}{2}$-approximation for continuous submodular maximization [18]. Although SGA is only guaranteed to find a local optimum, each iteration applies purely differentiable computations (a gradient step and projection onto the set $conv(\mathcal{X})$), and so the final output after T iterations will be differentiable as well. Provided that T is sufficiently large, this output will converge to a local optimum, which must satisfy the KKT conditions. Hence, we can apply our general approach to the local optimum returned by SGA. The following theorem shows that the local optima of the multilinear extension are differentiable:

Theorem 3. *Suppose that x^* is a local maximum of the multilinear extension, i.e,., $\nabla_x F(x^*, \theta) = 0$ and $\nabla_x^2 F(x^*, \theta) \succ 0$. Then, there exists a neighborhood $\mathcal{I}$ around x^* such that the maximizer of $F(\cdot, \theta)$ within $\mathcal{I} \cap conv(\mathcal{X})$ is differentiable almost everywhere as a function of θ, with $\frac{dx(\theta)}{d\theta}$ satisfying the conditions in Equation 2.*

We remark that Theorem 3 requires a local maximum, while gradient ascent may in theory find saddle points. However, recent work shows that random perturbations ensure that gradient ascent quickly escapes saddle points and finds an approximate local optimum [22].

Efficient backward pass: We now show how the terms needed to compute gradients via Equation 2 can be efficiently obtained. In particular, we need access to the optimal dual variable λ as well as the term $\frac{d\nabla_x F(x,\theta)}{d\theta}$. These were easy to obtain in the LP setting but the submodular setting requires some additional analysis. Nevertheless, we show that both can be obtained efficiently.

Optimal dual variables: SGA only produces the optimal primal variable x, not the corresponding dual variable λ which is required to solve Equation 2 in the backward pass. We show that for cardinality-constrained problems, we can obtain the optimal dual variables analytically given a primal solution x. Let λ_i^L be the dual variable associated with the constraint $x_i \geq 0$, λ_i^U with $x_i \leq 1$ and λ^S with $\sum_i x_i \leq k$. By differentiating the Lagrangian, any optimum satisfies

$$\nabla_{x_i} f(x) - \lambda_i^L + \lambda_i^U + \lambda_i^S = 0 \quad \forall i$$

where complementary slackness requires that $\lambda_i^L = 0$ if $x_i > 0$ and $\lambda_i^U = 0$ if $x_i < 1$. Further, it is easy to see that for all i with $0 < x_i < 1$, $\nabla_{x_i} f(x)$ must be equal. Otherwise, x could not be (locally) optimal since we could increase the objective by finding a pair i, j with $\nabla_{x_i} f(x) > \nabla_{x_j} f(x)$, increasing x_i, and decreasing x_j. Let ∇_* denote the shared gradient value for fractional entries. We can solve the above equation and express the optimal dual variables as

$$\lambda^S = -\nabla_*, \quad \lambda_i^L = \lambda^S - \nabla_{x_i} f, \quad \lambda_i^U = \nabla_{x_i} f - \lambda^S$$

where the expressions for λ_i^L and λ_i^U apply only when $x_i = 0$ and $x_i = 1$ respectively (otherwise, complementary slackness requires these variables be set to 0).

Computing $\mathbf{\frac{d}{d\theta}\nabla_x F(x,\theta)}$: We show that this term can be obtained in closed form for the case of probabilistic coverage functions, which includes many cases of practical interest (e.g. budget allocation, sensor placement, facility location, etc.). However, our framework can be applied to arbitrary submodular functions; we focus here on coverage functions just because they are particularly common in applications. A coverage function takes the following form. There a set of items U, and each $j \in U$ has a weight w_j. The algorithm can choose from a ground set V of actions. Each action a_i covers each item j independently with probability θ_{ij}. We consider the case where the probabilities θ are be unknown and must be predicted from data. For such problems, the multilinear extension has a closed form

$$F(x,\theta) = \sum_{j \in U} w_j \left(1 - \prod_{i \in V} 1 - x_{ij}\theta_{ij}\right)$$

and we can obtain the expression

$$\frac{d}{d\theta_{kj}}\nabla_{x_i}F(x,\theta) = \begin{cases} -\theta_{ij}x_k \prod_{\ell \neq i,k} 1 - x_\ell\theta_{\ell j} & \text{if } k \neq i \\ \prod_{k\neq i} 1 - x_k\theta_{kj} & \text{otherwise.} \end{cases}$$

5 Application: TB treatment

As discussed earlier, tuberculosis (TB) is a critical societal challenge. While many countries (including India) provide treatment free of charge, there are still many barriers to completing the full six-month antibiotic course required for first-line treatment. Treatment failures increase the risk of the continued presence of latent TB, as well as the development of drug-resistant strains. Accordingly, continued and proactive intervention by health workers is a critical part of the treatment process since the health worker can to act early to resolve potential adherence issues. However, TB workers are often severely resource-limited; e.g., a worker in India may be responsible for tens or hundreds of patients. Ideally, workers would make house visits to each at-risk patient, but this case load requires workers to prioritize which patients are at greatest risk.

We focus on a specific optimization problem that models the allocation of health workers to intervene with patients who are at risk in the near future. This prospective intervention is enabled by our real-time risk predictions and serves as an example of how our system can enable proactive, targeted action by providers. However, we emphasize that our system can be easily modified to capture other intervention problems. Such flexibility is one benefit to our technical approach, which allows the ML model to *automatically* adapt to the problem specified by a domain expert.

Our optimization problem models a health worker who plans a series of interventions 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*; details about why this is an appropriate objective can be found in [28]. Roughly, it 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 c_{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:

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

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.

The machine learning task is to predict the values of the c_{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 a set of fully connected layers. We refer to the final model as DeepNet; see [28] for details.

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. 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 (as they would benefit most from such a system). We included all patients, including those with no missed doses in the last week, since the overall resource allocation problem over locations must still account for them.

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.

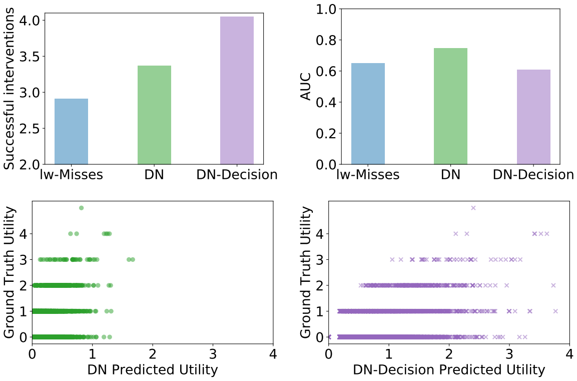

Fig. 1. Results for decision focused learning problem. Top row: successful interventions and AUC for each method. Bottom row: visualizations of model predictions.

Acknowledgements: We thank the Everwell team, especially Brandon Liu, Priyanka Ivatury, Amy Chen, and Bill Thies for sharing the data and thoughtfully advising us each step of the way. This work was supported by the Army Research Office (MURI W911NF1810208) and a National Science Foundation Graduate Research Fellowship.

References

1. 99DOTS: 99dots, https://www.99dots.org/
2. Amos, B., Kolter, J.Z.: Optnet: Differentiable optimization as a layer in neural networks. In: ICML (2017)
3. Belanger, D., Yang, B., McCallum, A.: End-to-end learning for structured prediction energy networks. In: ICML (2017)
4. Bertsimas, D., Dunn, J.: Optimal classification trees. Machine Learning **106**(7), 1039–1082 (2017)

5. Beygelzimer, A., Langford, J.: The offset tree for learning with partial labels. In: KDD (2009)
6. Bian, A.A., Mirzasoleiman, B., Buhmann, J.M., Krause, A.: Guaranteed non-convex optimization: Submodular maximization over continuous domains. In: AISTATS (2017)
7. Calinescu, G., Chekuri, C., Pál, M., Vondrák, J.: Maximizing a monotone submodular function subject to a matroid constraint. SIAM Journal on Computing **40**(6), 1740–1766 (2011)
8. Cook, P.F., Schmiege, S.J., Starr, W., Carrington, J.M., Bradley-Springer, L.: Prospective state and trait predictors of daily medication adherence behavior in hiv. Nursing research **66**(4), 275–285 (2017)
9. Djolonga, J., Krause, A.: Differentiable learning of submodular models. In: NIPS. pp. 1013–1023 (2017)
10. Domke, J.: Generic methods for optimization-based modeling. In: Artificial Intelligence and Statistics. pp. 318–326 (2012)
11. Donti, P., Amos, B., Kolter, J.Z.: Task-based end-to-end model learning in stochastic optimization. In: Advances in Neural Information Processing Systems. pp. 5484–5494 (2017)
12. Everwell: Everwell, http://www.everwell.org/
13. Fang, F., Nguyen, T.H., Pickles, R., Lam, W.Y., Clements, G.R., An, B., Singh, A., Tambe, M., Lemieux, A., et al.: Deploying paws: Field optimization of the protection assistant for wildlife security. In: IAAI. pp. 3966–3973 (2016)
14. Ford, B., Nguyen, T., Tambe, M., Sintov, N., Delle Fave, F.: Beware the soothsayer: From attack prediction accuracy to predictive reliability in security games. In: International Conference on Decision and Game Theory for Security (2015)
15. 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)
16. Gould, S., Fernando, B., Cherian, A., Anderson, P., Cruz, R.S., Guo, E.: On differentiating parameterized argmin and argmax problems with application to bi-level optimization. arXiv preprint arXiv:1607.05447 (2016)
17. 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)
18. Hassani, H., Soltanolkotabi, M., Karbasi, A.: Gradient Methods for Submodular Maximization. In: NIPS (2017)
19. Horvitz, E.: From data to predictions and decisions: Enabling evidence-based healthcare. Computing Community Consortium **6** (2010)
20. Horvitz, E., Mitchell, T.: From data to knowledge to action: A global enabler for the 21st century. Computing Community Consortium **1** (2010)
21. Iyer, R.K., Jegelka, S., Bilmes, J.A.: Monotone closure of relaxed constraints in submodular optimization. In: UAI (2014), https://dslpitt.org/uai/displayArticleDetails.jsp?mmnu=1smnu=2article$_id = 2471proceeding_id = 30$
22. Jin, C., Ge, R., Netrapalli, P., Kakade, S.M., Jordan, M.I.: How to escape saddle points efficiently. In: ICML (2017)

23. Kalhori, S.R.N., 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)
24. Kardas, P., Lewek, P., Matyjaszczyk, M.: Determinants of patient adherence: a review of systematic reviews. Frontiers in pharmacology **4**, 91 (2013)
25. Kempe, D., Kleinberg, J., Tardos, É.: Maximizing the Spread of Influence Through a Social Network. In: KDD (2003)
26. Khalil, E.B., Dai, H., Zhang, Y., Dilkina, B., Song, L.: Learning combinatorial optimization algorithms over graphs. In: NIPS (2017)
27. Khalil, E.B., Dilkina, B., Nemhauser, G.L., Ahmed, S., Shao, Y.: Learning to run heuristics in tree search. In: IJCAI (2017)
28. Killian, J.A., Wilder, B., Sharma, A., Choudhary, V., Dilkina, B., Tambe, M.: Learning to prescribe interventions for tuberculosis patients using digital adherence data. arXiv preprint arXiv:1902.01506 (2019)
29. Kim, K., Kim, B., Chung, A.J., Kwon, K., Choi, E., Nah, J.w.: Algorithm and system for improving the medication adherence of tuberculosis patients. In: 2018 International Conference on Information and Communication Technology Convergence (ICTC). pp. 914–916. IEEE (2018)
30. Kliiman, K., 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)
31. Korte, B., Vygen, J., Korte, B., Vygen, J.: Combinatorial optimization, vol. 2. Springer (2012)
32. 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)
33. Mukhopadhyay, A., Vorobeychik, Y., Dubey, A., Biswas, G.: Prioritized allocation of emergency responders based on a continuous-time incident prediction model. In: AAMAS. pp. 168–177 (2017)
34. Niculae, V., Martins, A.F., Blondel, M., Cardie, C.: Sparsemap: Differentiable sparse structured inference. In: ICML (2018)
35. Organization, W.H.: Global tuberculosis report 2018. World Health Organization (2018), licence: CC BY-NC-SA 3.0 IGO.
36. Organization, W.H., et al.: Handbook for the use of digital technologies to support tuberculosis medication adherence (2017)
37. Pellowski, J.A., Kalichman, S.C., Kalichman, M.O., Cherry, C.: Alcohol-antiretroviral therapy interactive toxicity beliefs and daily medication adherence and alcohol use among people living with hiv. AIDS care **28**(8), 963–970 (2016)
38. Platt, A.B., Localio, A.R., Brensinger, C.M., Cruess, D.G., Christie, J.D., Gross, R., Parker, C.S., Price, M., Metlay, J.P., Cohen, A., et al.: Can we predict daily adherence to warfarin?: Results from the international normalized ratio adherence and genetics (in-range) study. Chest **137**(4), 883–889 (2010)
39. Roy, N., Basu, M., Das, S., Mandal, A., Dutt, D., 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)
40. Shargie, E.B., Lindtjørn, B.: Determinants of treatment adherence among smear-positive pulmonary tuberculosis patients in southern ethiopia. PLoS medicine **4**(2), e37 (2007)

41. Subbaraman, R., de Mondesert, L., Musiimenta, A., Pai, M., Mayer, K.H., Thomas, B.E., Haberer, J.: Digital adherence technologies for the management of tuberculosis therapy: mapping the landscape and research priorities. BMJ global health **3**(5), e001018 (2018)
42. Thomas, A., Gopi, P., Santha, T., Chandrasekaran, V., Subramani, R., Selvakumar, N., Eusuff, S., Sadacharam, K., 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)
43. Tschiatschek, S., Sahin, A., Krause, A.: Differentiable submodular maximization. In: IJCAI (2018)
44. Tu, L., Gimpel, K.: Learning approximate inference networks for structured prediction. In: ICLR (2018)
45. Viappiani, P., Boutilier, C.: Optimal bayesian recommendation sets and myopically optimal choice query sets. In: NIPS (2010)
46. Vinyals, O., Fortunato, M., Jaitly, N.: Pointer networks. In: NIPS (2015)
47. Wang, H., Xie, H., Qiu, L., Yang, Y.R., Zhang, Y., Greenberg, A.: Cope: traffic engineering in dynamic networks. In: Sigcomm. vol. 6, p. 194 (2006)
48. Wilder, B., Dilkina, B., Tambe, M.: Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization. In: AAAI Conference on Artificial Intelligence (2018)
49. Xue, Y., Davies, I., Fink, D., Wood, C., Gomes, C.P.: Avicaching: A two stage game for bias reduction in citizen science. In: AAMAS. pp. 776–785 (2016)