# From Predictions to Decisions: Integrating Learning and Optimization for Wildlife Conservation Patrols

**Rachel Guo*** , **Lily Xu** and **Milind Tambe**
Harvard University
rguo@college.harvard.edu, lily_xu@g.harvard.edu, milind_tambe@harvard.edu

## Abstract

The illegal poaching of wildlife is a pressing conservation issue, driving the loss of biodiversity worldwide. To aid resource-constrained rangers on their search for vicious snares across expansive national parks, machine learning can be harnessed to predict the poaching hot-spots in a given protected area. However, beyond predictions on the relative risk of poaching activity in different areas of a park, rangers ultimately need guidance on selecting the best areas to patrol to maximize the number of animal traps they find en route given a budget constrained by limited patrolling resources. Traditionally, such a feat can be achieved through a two-stage approach, where a machine learning model is first trained to render predictions that are optimized upon thereafter. In this paper we introduce the the notion of combining the learning and combinatorial optimization process through a decision-focused learning approach for anti-poaching patrols. We find that we enhance the decision quality, as evidenced by an increase in the expected number of snares found by the targets we select for patrol using decision-focused learning compared to a two-stage approach. Thus, we improve the chances that rangers actually encounter snares in the target areas selected through our approach, preventing the loss of wildlife and biodiversity.

## 1 Introduction

Over past decades, illegal poaching of wildlife in protected areas has driven species to endangerment, accelerating the loss of biodiversity globally, with repercussions to life on Earth as grave as climate change at large [Pires and Moreto, 2016].

To combat illegal wildlife poaching, rangers are sent to patrol expansive national parks of thousands of kilometers squared and remove snares (Fig. 1). However, there are simply not enough patrol rangers to cover all areas of the park. To guide rangers to the most important areas to patrol, the Protection Assistant for Wildlife Security (PAWS) uses machine learning to predict the risk of poaching in different areas of the park based on geospatial and temporal features as well as historical patrol data [Xu *et al.*, 2020]. The system has been deployed to the SMART consortium (Spatial Monitoring and Reporting Tool), used by rangers on the ground in hundreds of national parks.

Figure 1: Rangers patrolling expansive areas in Queen Elizabeth National Park and removing snares. Photo courtesy of *Futurity* and *Wildlife Conservation Network*.

PAWS currently provides a risk map to rangers based on predictions made using a standard machine learning model, such as random forests and Gaussian processes (Fig. 2). The next step to helping rangers make sense of the predictions would be to include optimization to guide rangers to specific areas that would maximize the number of animal traps they find based on the predictions.

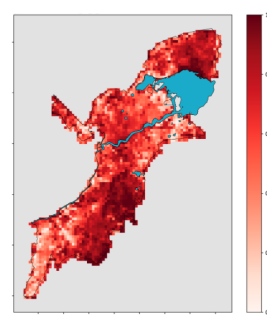


Figure 2: Risk map provided to rangers at Queen Elizabeth National Park in Uganda.

A natural way to optimize upon predictions is a traditional two-stage learning and optimization approach, where the machine learning model outputs predictions that are sent into the optimization algorithm, which selects the best targets to patrol given a budget constraint on patrolling resources. However, such an approach prioritizes achieving high predictive accuracy by minimizing loss during the training stage.

Instead, we propose a decision-focused learning approach that integrates the learning and optimization stages in a way

*Contact Author

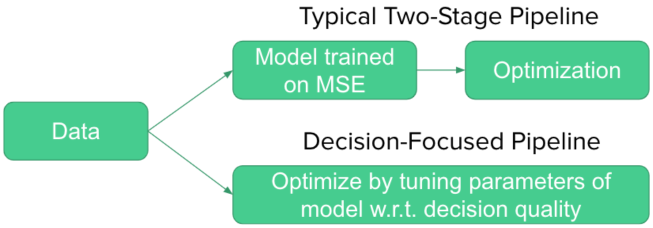
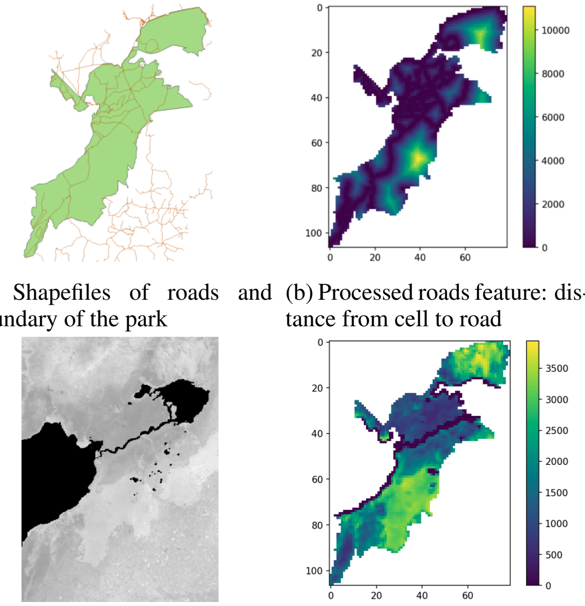



Figure 3: Comparison of two-stage and decision-focused pipelines.



(a) Shapefiles of roads and boundary of the park

(b) Processed roads feature: distance from cell to road

(c) Raw satellite imagery from GEE for NPP

(d) Processed NPP feature

Figure 4: We preprocess shapefiles for features such as roads (a) by computing the distance from each cell to the nearest road (b), as we believe distance to roads is predictive of poaching activity. We preprocess raster satellite imagery from GEE (c) such as net primary product (NPP) by interpolating the raster values to match our resolution size of $1 \times 1$ km cells (d).

that prioritizes the ultimate goal of the task at hand: to maximize the number of snares found through the patrol strategy our model prescribes (Fig. 3). In doing so, we find that, for every instance of 10 target cells, we increase the number of snares that rangers find by 0.021 in expectation given a budget of 3 out of the 10 target cells to patrol, compared to a traditional two-stage learning and optimization approach.

With our decision-focused model, we increase the chances that rangers actually find snares to remove in the cells we prescribe for them to patrol, with great potential for expediting anti-poaching efforts and slowing the loss of biodiversity.

## 2 Related Work

One of the major technical challenges of decision-focused learning is its reliance on computationally intensive differentiation over the decision space.

[Amos and Kolter, 2019] propose a differentiable optimization network architecture within a larger neural network structure that benefits from fast GPU computation. While the authors exemplify their approach through a $4\times4$ Sudoku grid, this form of optimization is inherently different from using such a technique on real-world data.

[Donti *et al.*, 2019] introduces the notion of training models to capture the ultimate task-based objective for which they will be used. However, these authors only consider synthetic domains such as the classic inventory stock problem. Moreover, their techniques are applied to stochastic programming, which is not the setting we consider in this paper.

[Wilder *et al.*, 2018] introduces integrating combinatorial optimization into the predictive process, which is the setting we consider as we aim to optimize for the top $k$ targets over a discrete set of targets given a budget $k$. However, [Wilder *et al.*, 2018] demonstrates its efficacy only over synthetic domains such as budget allocation of synthetic ad data.

In sum, little is known regarding the robustness of how decision-focused learning fares in real-world applications such as anti-poaching patrols, where data is incredibly class-imbalanced and incomplete.

## 3 Methods

### 3.1 Data Preprocessing

We use data from Queen Elizabeth National Park in Uganda from 2010 to 2016. Given a boundary of the park, we discretize the park region into $1 \times 1$ km cells. We similiarly discretize all geospatial and temporal features from the park into $1 \times 1$ km cells.

Some geospatial features such as road shapefiles and animal density TIF files come from the park managers themselves. Other geospatial features like landcover, elevation, and rivers as well as dynamic temporal data such as temperture and net primary product were automatically extracted from Google Earth Engine (GEE) remote sensing imagery [Guo *et al.*, 2020]. An example of preprocessing such data is shown in Fig. 4.

We also use historical patrol data given by the park manager that includes patrol waypoints including the time and location a cell was patrolled, and a binary indication of whether illegal poaching activity was found in the cell.

### 3.2 Problem Setup

The combinatorial optimization problem we seek to solve is of the form $\max_{z\in\mathcal{Z}} f(z,\theta)$, where $z$ is our decision variable, a binary vector of length $n$, and $\mathcal{Z} \subseteq \{0,1\}^n$, where $\mathcal{Z}$ is a discrete set of possible decisions. $f$ is our objective, which depends on the unknown parameter $\theta \in \Theta$, which is learned from the data.

For each training instance, we predict $\hat{\theta}$ from the feature vector $x \in \mathcal{X}$. Then we solve the optimization problem $\max_z f(z,\hat{\theta})$ to obtain the decision $z^*$. We thus define $z^*(\theta) = \arg\max_{z\in\mathcal{Z}} f(z,\hat{\theta})$. We seek to find a model $m : \mathcal{X} \rightarrow \Theta$ that maximizes

$$\mathbb{E}\left[f(z^*(m(x)),\theta)\right]$$

The traditional two-stage approach to first to this problem is to first train the model on a loss function, and to then use the trained model on optimization. However, such an approach prioritizes prediction accuracy rather than the decision task. Therefore, by combining the learning and optimization process, we train the model to actually make good decisions.

### 3.3 Combining Training and Optimization

We aim to integrate combinatorial optimization with the training process. We do so by directly training the model on the objective $\mathbb{E}\left[f(z^*(m(x)),\theta)\right]$. Doing so is difficult as the differentiation involves the $z^*(m(x))$ term. In particular, two major challenges entail dealing with the nondifferentiability of the discrete $z^*$ as well as differentiating through the argmax operation. Thus, we impose a continuous relaxation on the combinatorial optimization task, as [Wilder *et al.*, 2018] showed we can analytically obtain derivatives of the continuous optimizer with respect to the model's parameters. This allows us to train our model using a continuous form of the objective, and approximate the solution to the discrete problem by rounding the continuous points.

### 3.4 Continuous relaxation

We relax our constraint $z \in \mathcal{Z}$ to a continuous one: $z \in$ *conv*$(\mathcal{Z})$, with *conv* denoting the convex hull. When training our model, we compute gradients for the objective given by $\mathbb{E}\left[f(z^*(m(x)),\theta)\right]$, with $z^*$ replaced by its continuous counterpart. For a given training example, we derive the stochastic gradient estimate with respect to the model's parameters $w$ as

$$\frac{df(z(\hat{\theta}),\theta)}{dw} = \frac{df(z(\hat{\theta}),\theta)}{dz(\hat{\theta})}\frac{dz(\hat{\theta})}{d\hat{\theta}}\frac{d\hat{\theta}}{dw}$$

The first and third term are straightforward to compute; the challenge lies in computing the middle term, which tells us how the decision changes with respect to the model's predictions $\hat{\theta}$.

We solve this combinatorial optimization problem by using such a continuous relaxation and differentiating over the KKT conditions to solve for $z(\theta) = \arg\max_{z \in conv(\mathcal{Z})} f(z, \hat{\theta})$, the optimal solution for the continuous problem.

### 3.5 Experiments

After we preprocess the QENP data, we have 24,320 $1 \times 1$ km target cells. We divide 24,320 cells into instances of 10 cells such that we have 2,432 park instances. The size of each park instance is $10 \times 51$, covering 10 consecutive target cells with 34 geospatial features including elevation, temperature, roads, rivers, and 51 total features after one-hot encoding categorical features such as drainage direction and land cover classification. We used an 80-20 split for train and test.

Given budget $k$, we select the top $k$ targets out of 10 targets for each park instance. Using the decision-focused learning approach, we differentiate with respect to the decision i.e. the combinatorial selection of $k$ targets for a given park instance.

We run experiments for the decision-focused model, the traditional two stage model where training and optimization are separate, the optimal selection of top $k$ targets based on the ground truth data, and a random selection of the top $k$ targets for each park instance.

Table 1: Decision quality

| | $k=3$ | $k=5$ |
|---|---|---|
| Decision-Focused | $0.367 \pm 0.018$ | $0.607 \pm 0.029$ |
| Two-Stage | $0.345 \pm 0.011$ | $0.586 \pm 0.015$ |
| Optimal | $1.090 \pm 0.014$ | $1.140 \pm 0.020$ |
| Random | $0.360 \pm 0.007$ | $0.574 \pm 0.011$ |

Table 2: Prediction performance (MSE)

| | $k=3$ | $k=5$ |
|---|---|---|
| Decision-Focused | 1422.10 | 1359.55 |
| Two-Stage | 0.125 | 0.112 |

For our model $m$, we used a neural network architecture in both the decision-focused learning and two-stage learning approach with 2 fully connected layers, 40 hidden layers, and a ReLU activation.

For each model, we averaged the results over 4 runs. While we ideally would like to run on more runs to average over, the decision-focused approach turned out to be computationally expensive for differentiation over the argmax. The decision-focused model was trained over 40 iterations.

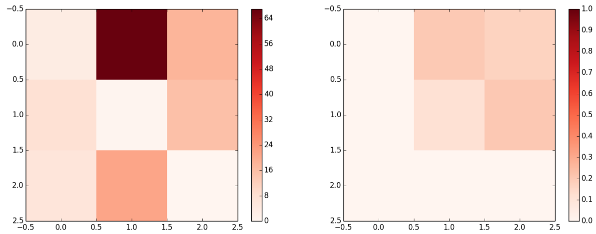
We tested different budget sizes $k = 3$ and $k = 5$ for park instances of 10 target cells. While we tried experimenting with larger park instances as well (and thus proportionally larger budget sizes) such as 25 target cells, computing the derivatives over the argmax with the continuous relaxation again proved computationally challenging.

## 4 Results

We assess performance by comparing the average number of cells, out of the top $k$ cells that we select for patrolling, that actually have snares.

Table 1 presents the average decision qualilty over four runs reported with the standard error of the mean. As seen in the results for the decision quality of our models in Table 1, our decision-focused model yields a better decision quality than the traditional two stage approach for both budgets $k = 3$ and $k = 5$. In particular, 0.021 more cells are found to have poaching using decision-focused learning compared to two-stage learning in expectation for budget of 3, and 0.022 more cells are found to have poaching using decision-focused learning compared to two-stage learning in expectation for budget of 5.

While this improvement may seem small, this is actually a drastic improvement considering the severe class imbalance in our ground truth poaching activity data, where we deal with only  10% of the data with positive instances. Indeed, we see that on average, Optimal selects finds 1.09 out of the top 3 cells chosen to patrol, suggesting that only around 1 of the 10 target cells have illegal activity in them on average. Moreover, increasing the budget by patrolling 2 additional cells only increases the number of snares that Optimal finds by 0.05, reinforcing the fact that a 0.055 increase in the number of snares found is a substantial result.




(a) Decision-Focused Predictions

(b) Two-Stage Predictions

Figure 5: Comparison of predictions made by the decision-focused model vs. the two-stage model on a single park instance of 9 target cells in a $3 \times 3$ km area.

While decision-focused learning performs better than Random, as expected, it is interesting to note that Random performs pretty well, surpassing Two-Stage for $k = 3$, considering that it randomly selects the top $k$ cells to patrol. We believe this phenomenon can be attributed to the fact that we are dealing with small park instances of size 10, a small budget of 3, and an incredibly class-imbalanced dataset, all of which combine to make it so that, by random chance alone, it is possible to select a decent set of cells that happen to encapsulate the one or two illegal activity instances for a given park instance.

As we see in Table 2, the mean squared errors (MSEs) are large for the decision-focused model as the optimizer does not consider MSE when trying to improve decision quality. MSE is not something we are concerned about in our evaluation, as our goal is to maximize our decision quality rather than MSE loss. However, we display them here to illustrate the phenomenon that decision-focused learning improves decision quality despite having lower prediction accuracy.

Note that the two plots in Fig. 5 are on different scales as we do not place constraints on the predictions for the decision-focused model; therefore, the relative values can be compared between the two plots. Meanwhile, the two-stage predictions are constrained between 0 and 1 as they represent the probability of finding a snare in a target cell.

Observe also that the decision-focused model generates richer predictions through the broad range of colors, and thus values, represented. On the other hand, the predictions from the two-stage model seem to all hover around a 0.2 probability of poaching. Moreover, five cells in the two-stage model were predicted to have zero probability of poaching, in comparison to two cells in the decision-focused model that were predicted to have zero probability of poaching, suggesting that the decision-focused model's predictions are more informative. Finally, we note the difference between the location of the cells found to have higher risk of poaching in the two models. For example, the bottom-middle cell in the decision-focused plot is predicted to have higher risk of poaching than the cell in the top-right corner, while the bottom-middle cell in the two-stage plot is predicted to have much lower probability of poaching than the cell in the top-right corner.

It is not immediately clear which plot's predictions are more accurate, though we can wager that the two-stage predictiions, while less varying in its prediction values, are more accurate to the ground truth considering its exceptionally small MSE. This will make for interesting future explorations as we seek to understand why the predictions between the two models differ in this spatial manner. Perhaps using convolutional neural networks will make good use of the spatial patterns available in our data to further enrich our predictions and make for closer comparisons between the decision-focused model and two-stage model. Such investigations can help us better understand how decision-focused learning makes its predictions, and why decision-focused learning performs better than the traditional two-stage model.

## 5 Conclusion

In this paper, we introduced a decision-focused approach in the detection of poaching activity. We find that decision-focused learning leads to a better decision quality than traditional two-stage training for maximizing the number of snares found during anti-poaching patrols. That is, a larger fraction of the top $k$ targets selected by the decision-focused model actually has poaching activity in expectation, and yielding a higher reward than two-stage training. Using our decision-focused approach, rangers are more likely to find snares in the cells that the decision-focused model selects for patrolling, which has major implications in the slowing of the loss of biodiversity to illegal wildlife poaching.

## 6 Future Work

In the future, we aim to use convolutional neural networks in tandem with our decision-focused approach to further improve upon our decision quality using the spatial patterns in our geospatial data. This will be interesting from a technical perspective as well, as decision-focused learning has not yet been applied to CNNs to our knowledge.

## 7 Ethics and Broader Impact

While our decision-focused approach improves the chances that rangers actually find snares in the cells they patrol, we note that there are ethical considerations we must address. In particular, our approach may make it more likely that rangers encounter actual poachers as they target the most likely cells to have snares. However, many of these poachers should not be considered criminals as they are merely villagers trying to survive and feed their family. The actual criminals are the larger illegal smuggling businesses that exploit these people to poach for their profit. Therefore, it will be increasingly important to work with local governments and agencies that use our system to ensure the way rangers and officials deal with poachers encountered en route is ethical and fair to their circumstances.

## Acknowledgments

We thank the Uganda Wildlife Authority for QENP data, as well as the SMART consortium for the partnership. We also thank the rangers doing critical work on the ground to protect our wildlife.

## References


[Amos and Kolter, 2019] Brandon Amos and J. Zico Kolter. Optnet: Differentiable optimization as a layer in neural networks, 2019.

[Donti *et al.*, 2019] Priya L. Donti, Brandon Amos, and J. Zico Kolter. Task-based end-to-end model learning in stochastic optimization, 2019.

[Guo *et al.*, 2020] Rachel Guo, Lily Xu, Drew Cronin, Francis Okeke, Andrew Plumptre, and Milind Tambe. Enhancing poaching predictions for under-resourced wildlife conservation parks using remote sensing imagery, 2020.

[Pires and Moreto, 2016] S. F. Pires and W. Moreto. The illegal wildlife trade, 2016.

[Wilder *et al.*, 2018] Bryan Wilder, Bistra Dilkina, and Milind Tambe. Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization, 2018.

[Xu *et al.*, 2020] L. Xu, S. Gholami, S. McCarthy, B. Dilkina, A. Plumptre, M. Tambe, R. Singh, M. Nsubuga, J. Mabonga, M. Driciru, F. Wanyama, A. Rwetsiba, T. Okello, and E. Enyel. Stay ahead of poachers: Illegal wildlife poaching prediction and patrol planning under uncertainty with field test evaluations. In *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, pages 1898–1901, 2020.