Neur2BiLO: Neural Bilevel Optimization

Justin Dumouchelle • Esther Julien • Jannis Kurtz • Elias B. Khalil

University of Toronto · TU Delft · University of Amsterdam

Introduction

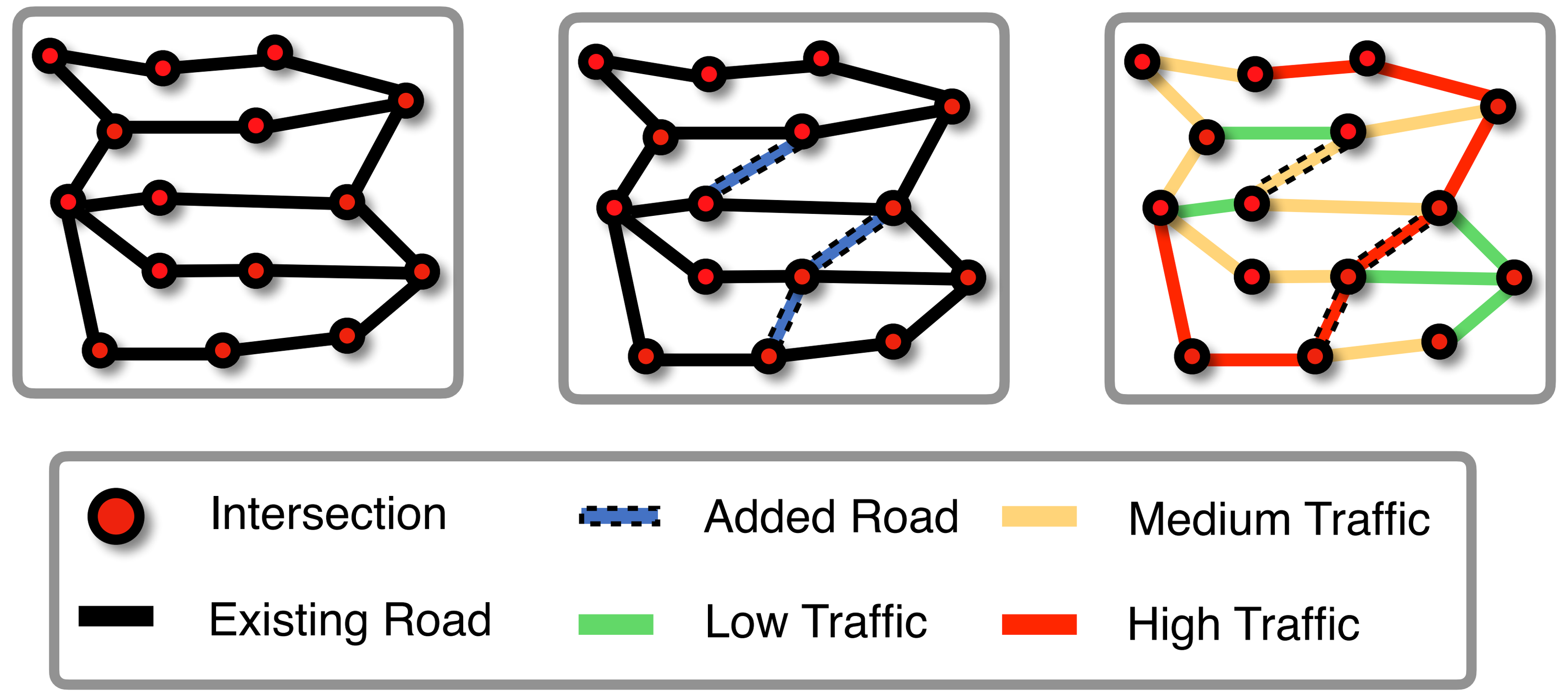
- ***Summary***: A fast learning-based heuristic for *mixed integer (non-)linear bilevel optimization*
- ***Objective***: Find the ***leader / upper-level solution*** $(\mathbf{x})$ that minimize the objective $F(\mathbf{x},\mathbf{y})$ while considering the ***follower / lower-level solution*** $(\mathbf{y})$ will maximize their own objective $f(\mathbf{x},\mathbf{y})$, subject to any constraints

$$\min_{\mathbf{x}\in\mathcal{X},\mathbf{y}} \; F(\mathbf{x},\mathbf{y})$$
$$\text{s.t.} \; G(\mathbf{x},\mathbf{y}) \geq \mathbf{0}$$
$$\mathbf{y} \in \arg\max_{\mathbf{y}'\in\mathcal{Y}} \{f(\mathbf{x},\mathbf{y}') : g(\mathbf{x},\mathbf{y}') \geq \mathbf{0}\}$$

- ***Challenge***: Integer decisions variables make gradient and duality-based methods inapplicable
- ***STOA***: Fast problem-specific heuristics for well-studied problems and less efficient general-purpose algorithms

Example

- ***Discrete Network Design Problem***: A city planner wants to build a subset of roads $(\mathbf{x})$ to a road network to minimize the total travel time in the network, $F(\mathbf{x},\mathbf{y})$. Based on the roads added, drivers will reach a user equilibria, $\mathbf{f}(\mathbf{x},\mathbf{y})$, i.e., a solution $(\mathbf{y})$ in which no user can unilaterally reduce their travel time



Methodology

1 Train a ML model to approximate the *value functions* of the leader and follower

Upper-Level Approximation (ULA)

$$\text{NN}^u_\Theta(\mathbf{x}) \approx F(\mathbf{x},\mathbf{y}^\star)$$

Lower-Level Approximation (LLA)

$$\text{NN}^l_\Theta(\mathbf{x}) \approx \max_{\mathbf{y}\in\mathcal{Y}}\{f(\mathbf{x},\mathbf{y}) : g(\mathbf{x},\mathbf{y}) \geq \mathbf{0}\}$$

2 Reformulate as a ***single-level mixed integer program*** by optimizing over predictive model

ULA

$$\min_{\mathbf{x}\in\mathcal{X}}\{\text{NN}^u_\Theta(\mathbf{x}) : G(\mathbf{x}) \geq \mathbf{0}\}$$

- Less variables (no $\mathbf{y}$)
- Faster to optimize
- Does not model coupling constraints $G(\mathbf{x},\mathbf{y})$
- May be infeasible

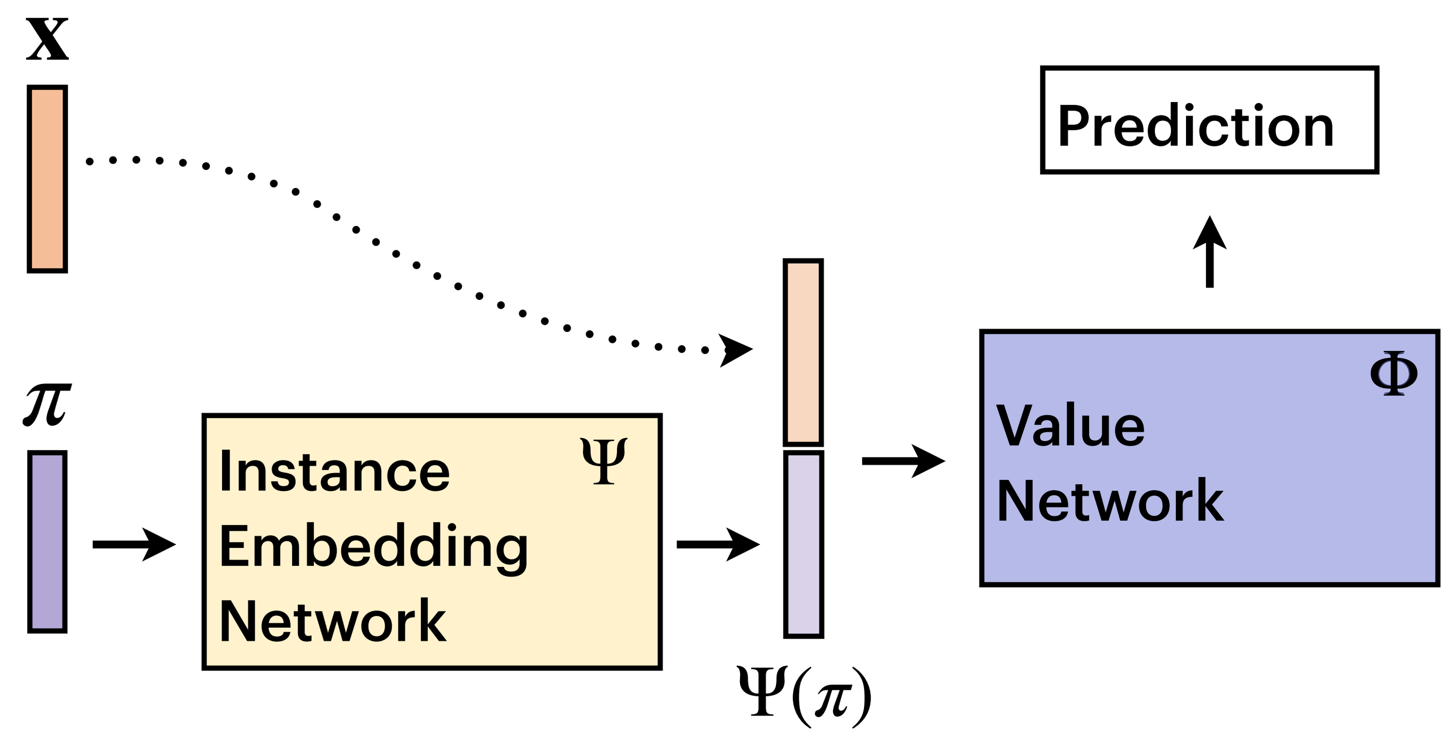
LLA

$$\min_{\mathbf{x}\in\mathcal{X},\mathbf{y}} \; F(\mathbf{x},\mathbf{y}) + \lambda s$$
$$\text{s.t.} \; G(\mathbf{x},\mathbf{y}) \geq \mathbf{0}$$
$$g(\mathbf{x},\mathbf{y}) \geq \mathbf{0}$$
$$f(\mathbf{x},\mathbf{y}) \geq \text{NN}^l_\Theta(\mathbf{x}) - s$$

- Requires lower-level variables $(\mathbf{y})$, a slack variable (s), and hyperparameter (λ)
- Always feasible
- Better for highly constrained problems

Architecture

- For ***data-driven settings*** where we want to train one model across variables instances ($\pi \in \Pi$)



- ***Set / graph-based architectures*** can be used to create models that handle invariance in ***variable ordering*** and allow generalization to varying ***instances sizes***

Guarantees

- For min max problems, i.e., $F = f$, Neur2BiLO computes a solution with an optimality gap bounded by an additive function of the prediction error

ULA

If $|\text{NN}^u_\Theta(\mathbf{x}) - \Phi(\mathbf{x})| \leq \alpha, \; \forall \mathbf{x}\in\mathcal{X}$, then the upper-level approximation returns a solution $(\mathbf{x}^\star,\mathbf{y}^\star)$ that is bounded by $f(\mathbf{x}^\star,\mathbf{y}^\star) \leq \textbf{opt} + 2\alpha$

LLA

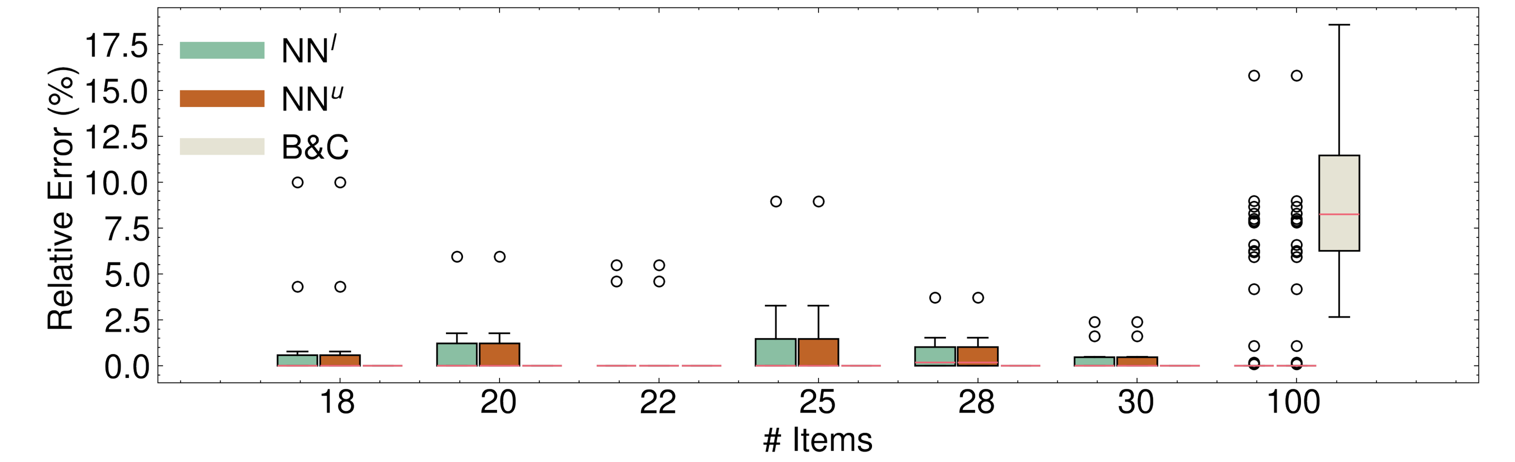
If $|\text{NN}^l_\Theta(\mathbf{x}) - \Phi(\mathbf{x})| \leq \alpha, \; \forall \mathbf{x}\in\mathcal{X}$, then the lower-level approximation returns a solution $(\mathbf{x}^\star,\mathbf{y}^\star)$ that is bounded by $f(\mathbf{x}^\star,\mathbf{y}^\star) \leq \textbf{opt} + 3\alpha + \frac{2}{\lambda}\Delta$, where

$$\Delta = \max\{f(\mathbf{x},\mathbf{y}) - f(\mathbf{x},\mathbf{y}'), \forall \mathbf{x}\in\mathcal{X}, \mathbf{y},\mathbf{y}'\in\mathcal{Y}\}$$

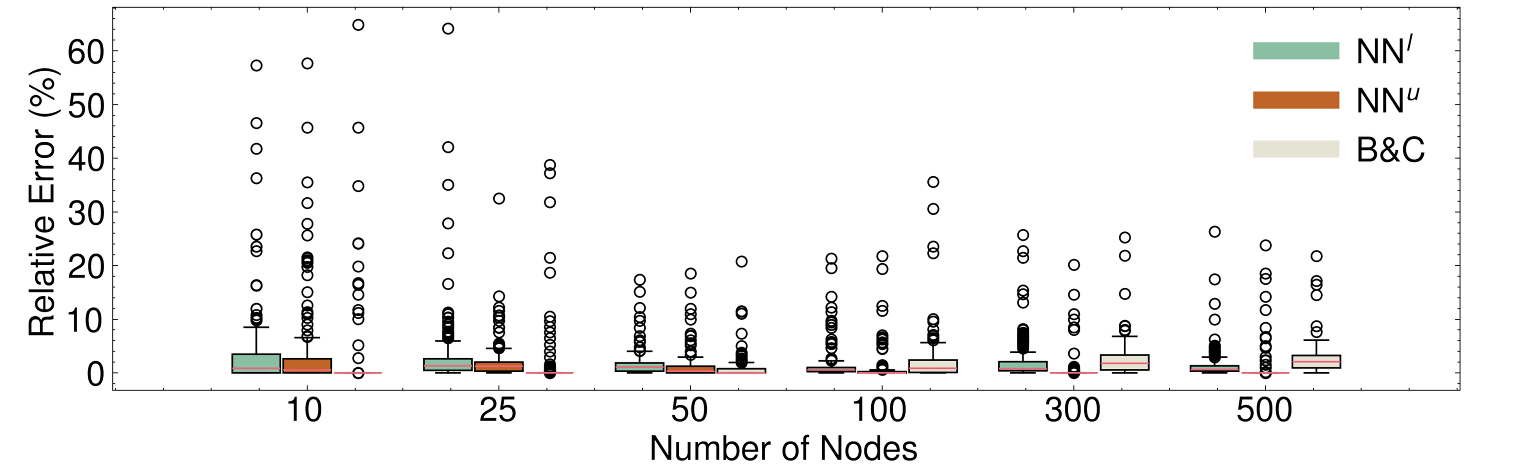
Computational Results

- ***Efficiency***: Neur2BiLO finds solutions 100-1000x faster
- ***Solution Quality***: Neur2BiLO finds STOA solutions on larger instances

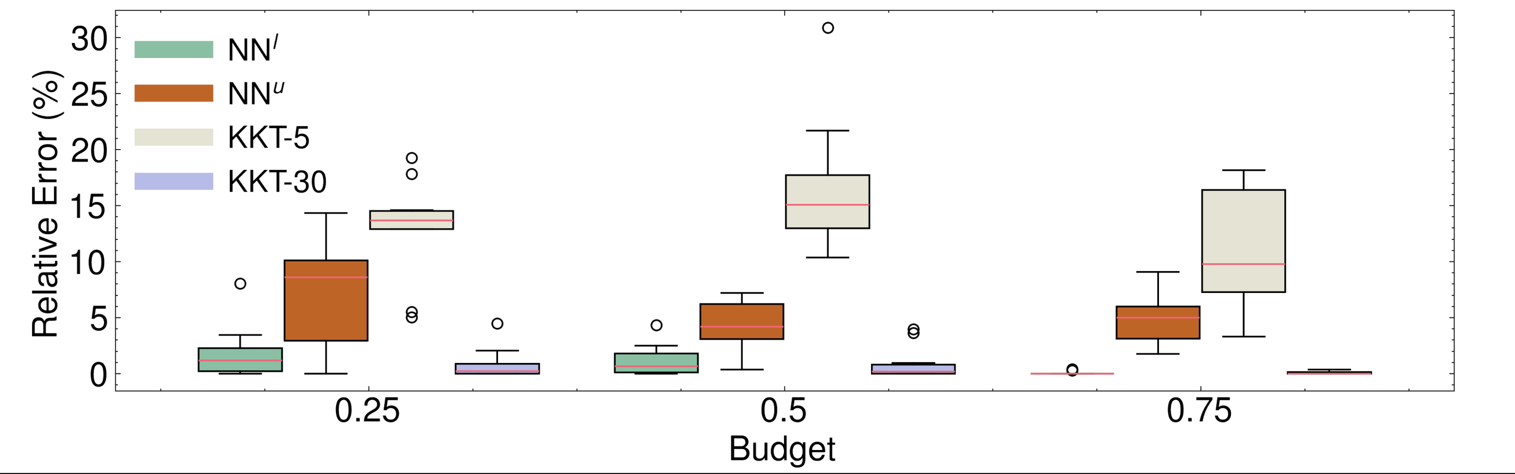
Knapsack Interdiction Problem *(Combinatorial Interdiction)*



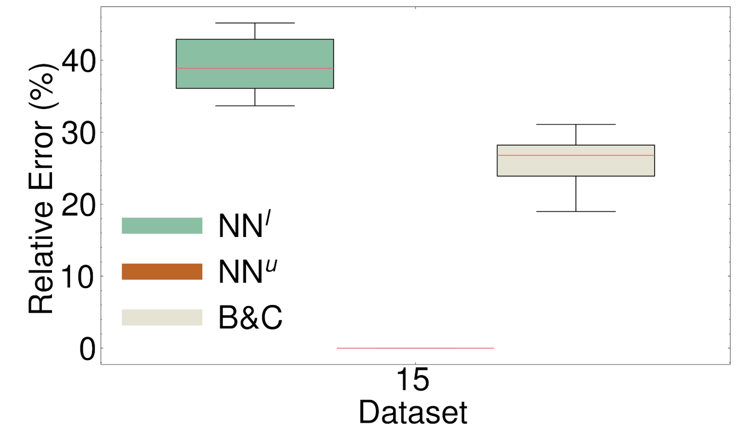
Critical Node Problem *(Security)*



Discrete Network Design Problem *(Transportation)*



Donor Recipient Problem *(Healthcare)*



Paper

Project Page