Optimization with Constraint Learning LAB

Part I

Speakers: Ilker Birbil & Donato Maragno
University of Amsterdam

February 8, 2023

Agenda

▪ Introduction to Constraint Learning

▪ Chemotherapy case study

▪ Embedding Decision trees

▪ Embedding Neural networks

▪ Trust region constraints

▪ Hands-on tutorial

Constraint Learning

University of Amsterdam IBM

Mixed-Integer Optimization (MIO)

Powerful tool that allows us to optimize a given objective subject to various constraints.

Many real-life optimization problems contain one or more constraints or objectives for which **there are no explicit formulae.**

Constraint Learning

Mixed-Integer Optimization (MIO)

Powerful tool that allows us to optimize a given objective subject to various constraints.

Many real-life optimization problems contain one or more constraints or objectives for which **there are no explicit formulae.**

Machine Learning (ML)

Data is available and machine learning models can be used to **learn the constraints.**

Constraint Learning

$\boldsymbol{x}$ Decision variables

$\boldsymbol{w}$ Contextual variables

$\boldsymbol{y}$ Predicted outcomes

Constraint Learning

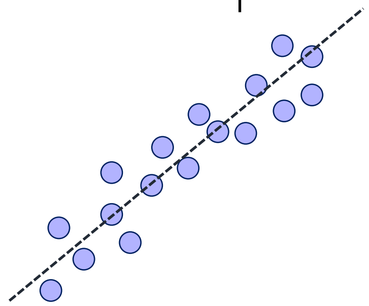
$\boldsymbol{x}$ Decision variables

$\boldsymbol{w}$ Contextual variables

$\boldsymbol{y}$ Predicted outcomes

$$\boldsymbol{y} = \hat{\boldsymbol{h}}_{\mathcal{D}}(\boldsymbol{x}, \boldsymbol{w})$$

$$D = \{(\overline{\boldsymbol{x}}_i, \overline{\boldsymbol{w}}_i, \overline{\boldsymbol{y}}_{\boldsymbol{i}})\}_{i=1}^{N} \longrightarrow$$



Constraint Learning

$$\min_{\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{y}\in\mathbb{R}^k} \quad f(\boldsymbol{x},\boldsymbol{w},\boldsymbol{y})$$

x Decision variables

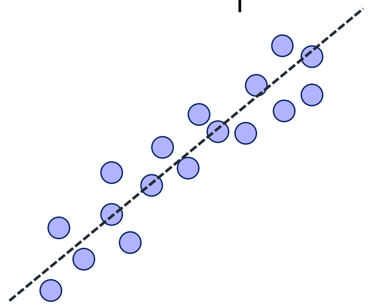
w Contextual variables

y Predicted outcomes

$$\text{s.t.} \quad \boldsymbol{g}(\boldsymbol{x},\boldsymbol{w},\boldsymbol{y}) \leq \boldsymbol{0}$$

$$\boxed{\boldsymbol{y} = \hat{\boldsymbol{h}}_{\mathcal{D}}(\boldsymbol{x},\boldsymbol{w})}$$

$$D = \{(\overline{\boldsymbol{x}}_i,\overline{\boldsymbol{w}}_i,\overline{\boldsymbol{y}}_i)\}_{i=1}^{N} \longrightarrow$$



Constraint Learning

$$\min_{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{y}\in\mathbb{R}^k} \quad f(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{y})$$

$\boldsymbol{x}$ Decision variables

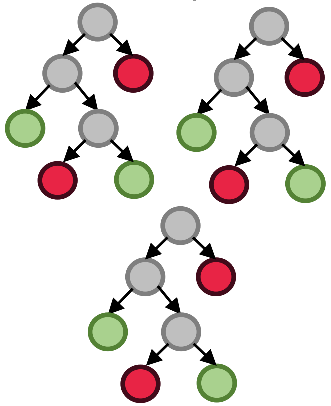
$\boldsymbol{w}$ Contextual variables

$\boldsymbol{y}$ Predicted outcomes

$$\text{s.t. } \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{y}) \leq \boldsymbol{0}$$

$$\boxed{\boldsymbol{y} = \hat{\boldsymbol{h}}_{\mathcal{D}}(\boldsymbol{x}, \boldsymbol{w})}$$

$$D = \{(\overline{\boldsymbol{x}}_i, \overline{\boldsymbol{w}}_i, \overline{\boldsymbol{y}}_i)\}_{i=1}^{N} \longrightarrow$$



Constraint Learning

$$\min_{\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{y}\in\mathbb{R}^k} f(\boldsymbol{x},\boldsymbol{w},\boldsymbol{y})$$

$\boldsymbol{x}$ Decision variables

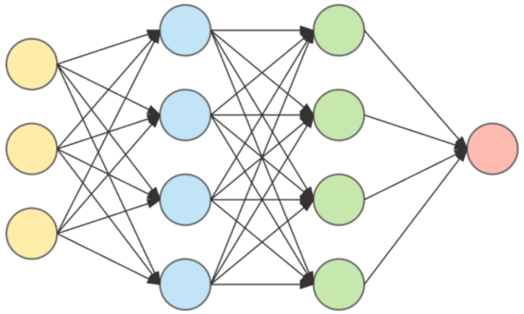
$\boldsymbol{w}$ Contextual variables

$\boldsymbol{y}$ Predicted outcomes

$$\text{s.t. } \boldsymbol{g}(\boldsymbol{x},\boldsymbol{w},\boldsymbol{y}) \leq \boldsymbol{0}$$

$$\boxed{\boldsymbol{y} = \hat{\boldsymbol{h}}_{\mathcal{D}}(\boldsymbol{x},\boldsymbol{w})}$$

$$D = \{(\overline{\boldsymbol{x}}_i, \overline{\boldsymbol{w}}_i, \overline{\boldsymbol{y}}_i)\}_{i=1}^{N} \longrightarrow$$



Chemotherapy Case Study

UNIVERSITY OF AMSTERDAM IBM

In this case study, we extend the work of Bertsimas et al. (2016)* in the design of chemotherapy regimens for advanced gastric cancer. Given a new study cohort and study characteristics, we would like to optimize a chemotherapy regimen to **maximize the cohort's survival subject to constraint on different types of toxicity**.

$$\boldsymbol{x}_b^d = \mathbb{I}(\text{drug } d \text{ is administered}),$$
$$\boldsymbol{x}_a^d = \text{average daily dose of drug } d,$$
$$\boldsymbol{x}_i^d = \text{maximum instantaneous dose of drug } d.$$

$$\begin{aligned}
\min_{\boldsymbol{x},\boldsymbol{y}} \quad & y_{OS} \\
\text{s.t.}\; & y_i \le \tau_i, && i \in \mathcal{Y}_C, \\
& y_i = \hat{h}_i(\boldsymbol{x}, \boldsymbol{w}), && i \in \mathcal{Y}_C, \\
& y_{OS} = \hat{h}_{OS}(\boldsymbol{x}, \boldsymbol{w}), \\
& \sum_d \boldsymbol{x}_b^d \le 3, \\
& \boldsymbol{x}_b \in \{0,1\}^d, \\
& \boldsymbol{x} \in \mathcal{X}(\boldsymbol{w}).
\end{aligned}$$

(*) Bertsimas D, O'Hair A, Relyea S, Silberholz J (2016) An analytics approach to designing combination chemotherapy regimens for cancer. Management Science 62(5):1511–1531, ISSN 15265501

Chemotherapy Case Study

In this case study, we extend the work of Bertsimas et al. (2016)* in the design of chemotherapy regimens for advanced gastric cancer. Given a new study cohort and study characteristics, we would like to optimize a chemotherapy regimen to **maximize the cohort's survival subject to constraint on different types of toxicity**.

$$\begin{aligned} x_b^d &= \mathbb{I}(\text{drug } d \text{ is administered}),\\ x_a^d &= \text{average daily dose of drug } d,\\ x_i^d &= \text{maximum instantaneous dose of drug } d. \end{aligned}$$

$$\begin{aligned} \min_{x,y}\ & y_{OS} \\ \text{s.t.}\ & y_i \le \tau_i, && i \in \mathcal{Y}_C,\\ & y_i = \hat{h}_i(x, w), && i \in \mathcal{Y}_C,\\ & y_{OS} = \hat{h}_{OS}(x, w),\\ & \sum_d x_b^d \le 3,\\ & x_b \in \{0,1\}^d,\\ & x \in \mathcal{X}(w). \end{aligned}$$

Cohort contextual variables

- Gender
- Age
- primary site breakdown
- ecog score
- …

(*) Bertsimas D, O'Hair A, Relyea S, Silberholz J (2016) An analytics approach to designing combination chemotherapy regimens for cancer. Management Science 62(5):1511–1531, ISSN 15265501

Chemotherapy Case Study

UNIVERSITY OF AMSTERDAM IBM

In this case study, we extend the work of [Bertsimas et al. (2016)]* in the design of chemotherapy regimens for advanced gastric cancer. Given a new study cohort and study characteristics, we would like to optimize a chemotherapy regimen to **maximize the cohort's survival subject to constraint on different types of toxicity**.

$\boldsymbol{x}_b^d = \mathbb{I}(\text{drug } d \text{ is administered}),$
$\boldsymbol{x}_a^d = \text{average daily dose of drug } d,$
$\boldsymbol{x}_i^d = \text{maximum instantaneous dose of drug } d.$

Overall Survival

$$
\begin{aligned}
\min_{\boldsymbol{x},\boldsymbol{y}} \quad & y_{OS} \\
\text{s.t.}\; & y_i \le \tau_i, && i \in \mathcal{Y}_C, \\
& y_i = \hat{h}_i(\boldsymbol{x}, \boldsymbol{w}), && i \in \mathcal{Y}_C, \\
& y_{OS} = \hat{h}_{OS}(\boldsymbol{x}, \boldsymbol{w}), \\
& \sum_d \boldsymbol{x}_b^d \le 3, \\
& \boldsymbol{x}_b \in \{0,1\}^d, \\
& \boldsymbol{x} \in \mathcal{X}(\boldsymbol{w}).
\end{aligned}
$$

Cohort contextual variables

- Gender
- Age
- primary site breakdown
- ecog score
- …

(*) Bertsimas D, O'Hair A, Relyea S, Silberholz J (2016) An analytics approach to designing combination chemotherapy regimens for cancer. Management Science 62(5):1511–1531, ISSN 15265501

Chemotherapy Case Study

In this case study, we extend the work of Bertsimas et al. (2016)* in the design of chemotherapy regimens for advanced gastric cancer. Given a new study cohort and study characteristics, we would like to optimize a chemotherapy regimen to **maximize the cohort's survival subject to constraint on different types of toxicity**.

$$
\begin{aligned}
\boldsymbol{x}_b^d &= \mathbb{I}(\text{drug } d \text{ is administered}),\\
\boldsymbol{x}_a^d &= \text{average daily dose of drug } d,\\
\boldsymbol{x}_i^d &= \text{maximum instantaneous dose of drug } d.
\end{aligned}
$$

Overall Survival

$$
\begin{aligned}
\min_{\boldsymbol{x},\boldsymbol{y}} \quad & y_{OS} \\
\text{s.t.}\; & y_i \le \tau_i, && i \in \mathcal{Y}_C,\\
& y_i = \hat{h}_i(\boldsymbol{x}, \boldsymbol{w}), && i \in \mathcal{Y}_C,\\
& y_{OS} = \hat{h}_{OS}(\boldsymbol{x}, \boldsymbol{w}),\\
& \sum_d \boldsymbol{x}_b^d \le 3,\\
& \boldsymbol{x}_b \in \{0,1\}^d,\\
& \boldsymbol{x} \in \mathcal{X}(\boldsymbol{w}).
\end{aligned}
$$

Toxicities

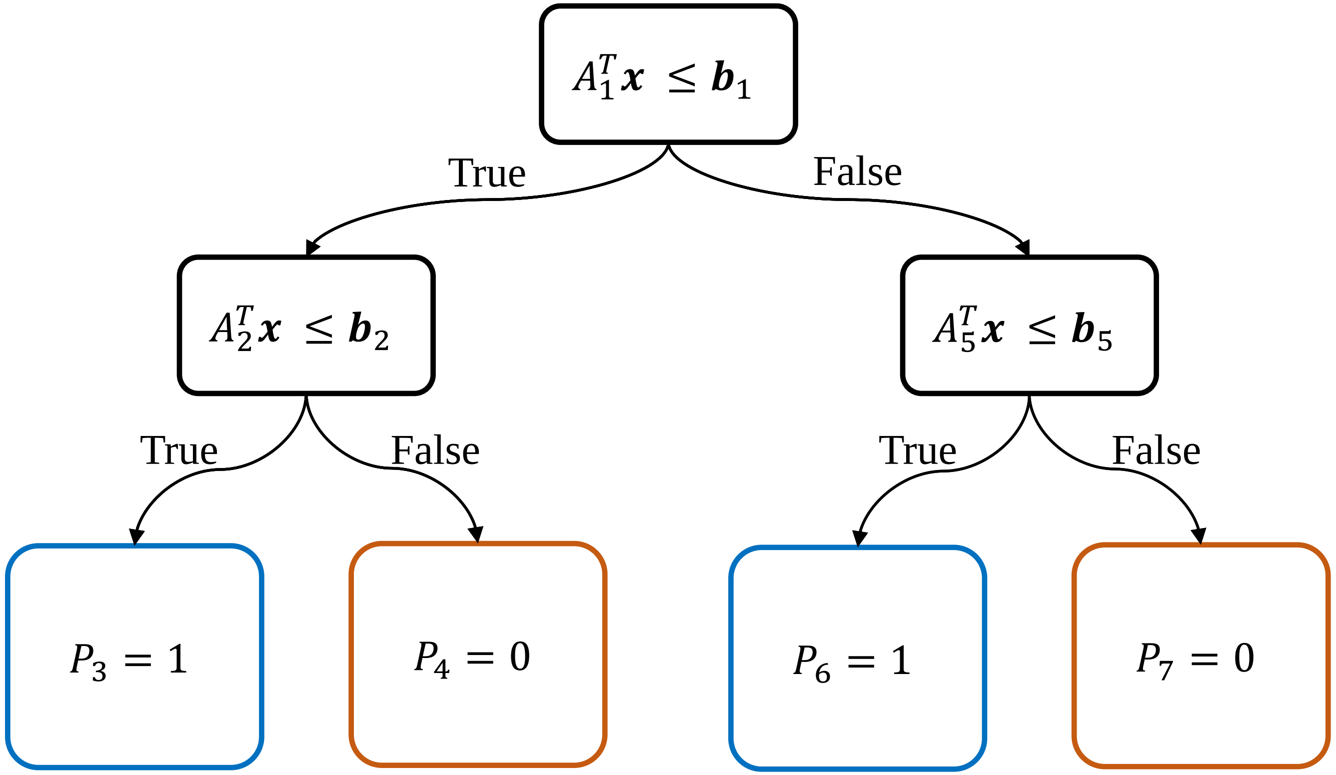
- Grade 3/4 constitutional
- Infection
- Neurological
- Grade 4 blood
- …

Cohort contextual variables

- Gender
- Age
- primary site breakdown
- ecog score
- …

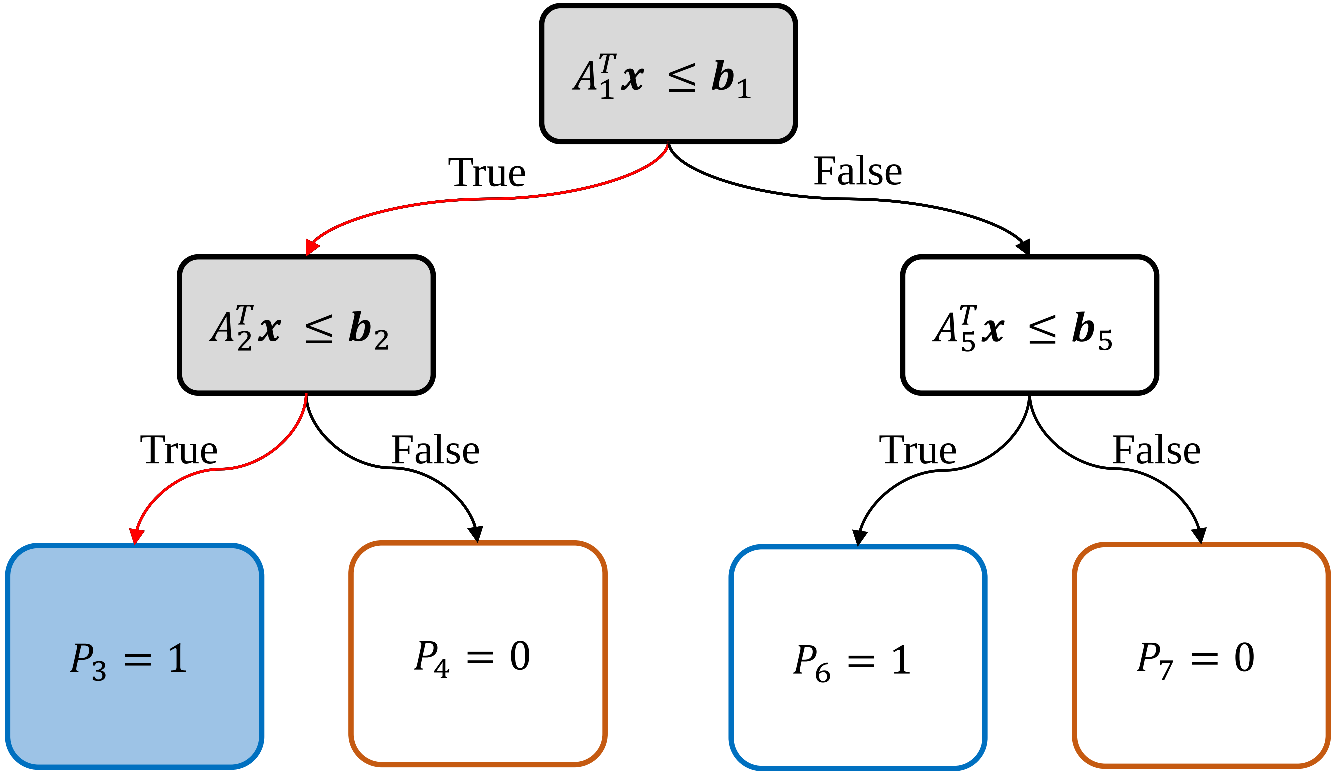
(*) Bertsimas D, O'Hair A, Relyea S, Silberholz J (2016) An analytics approach to designing combination chemotherapy regimens for cancer. Management Science 62(5):1511–1531, ISSN 15265501

Embedding Decision Trees



Embedding Decision Trees

University of Amsterdam IBM



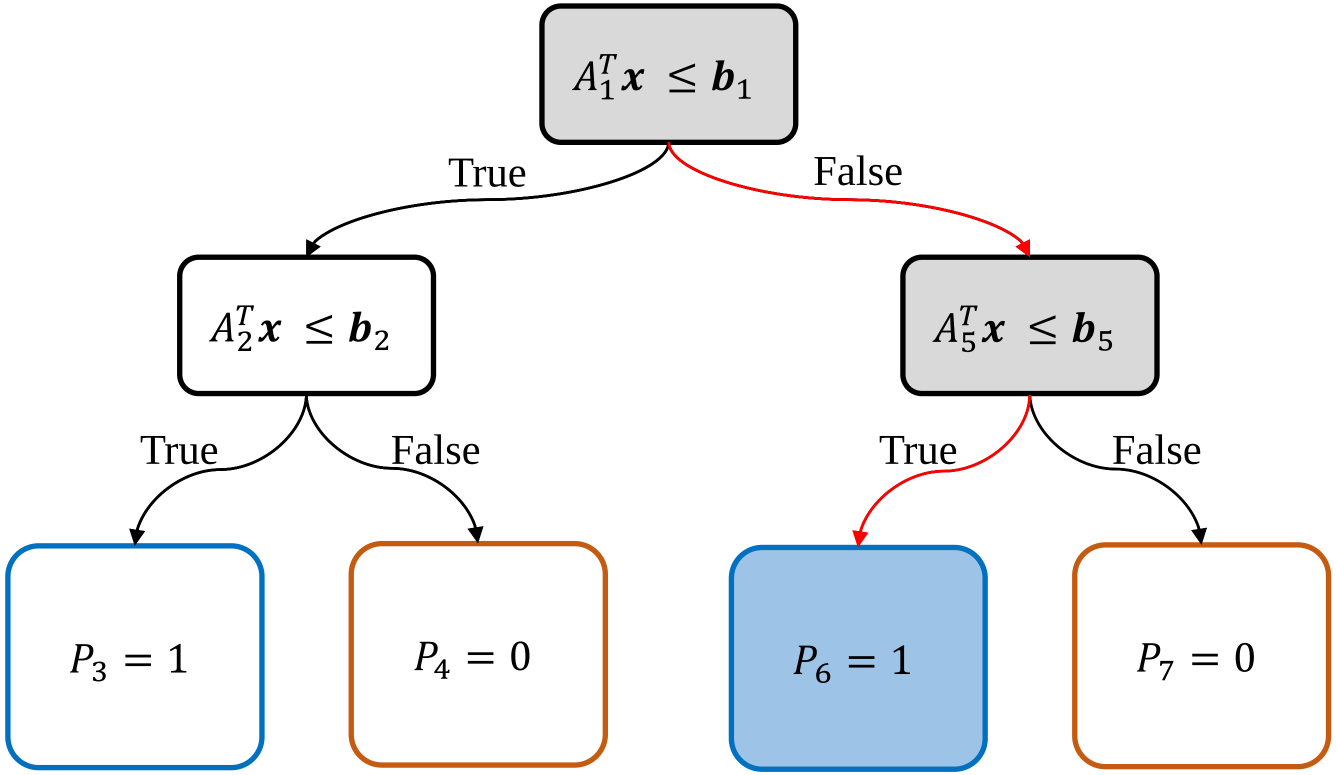
Conceptual model

$$\left\{ \begin{array}{l} \min_{\boldsymbol{x}} \; f(\boldsymbol{x}) \\ s.t. \; \boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0} \end{array} \right.$$

P_3

$$A_1^T\boldsymbol{x} \le \boldsymbol{b}_1$$
$$A_2^T\boldsymbol{x} \le \boldsymbol{b}_2$$

Embedding Decision Trees



Conceptual model

$$\begin{cases} \min\limits_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ s.t. & \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} \end{cases}$$

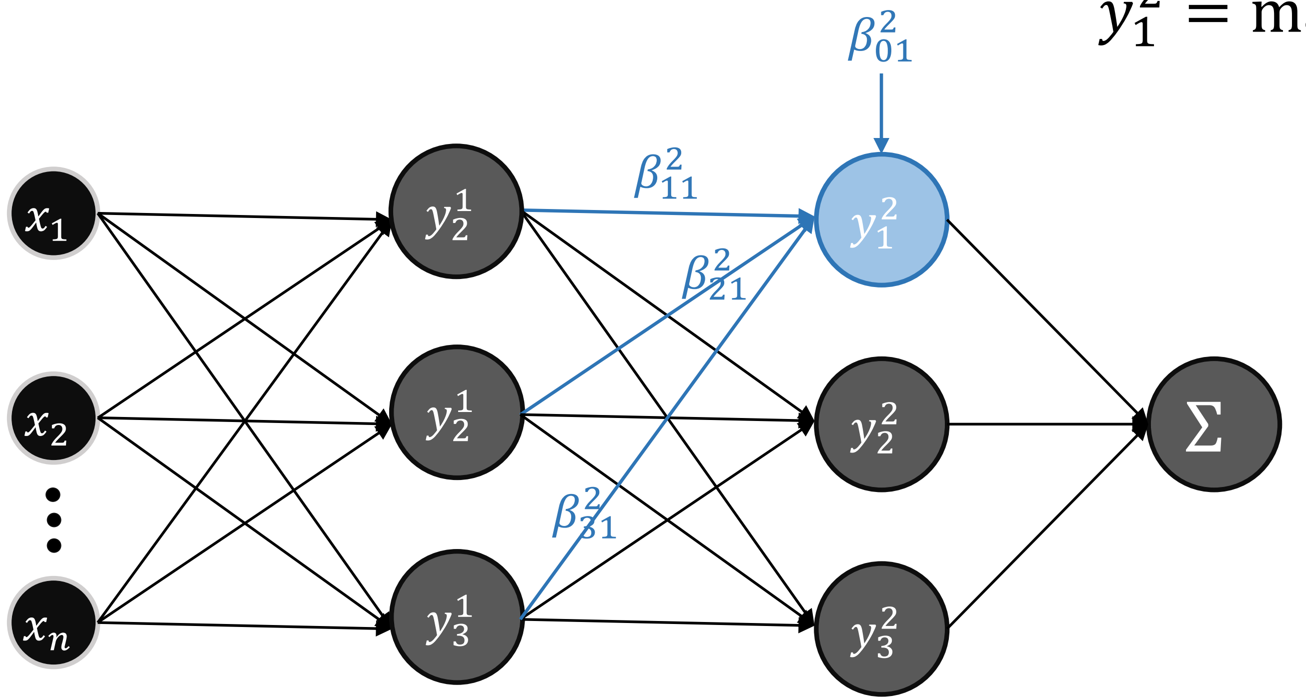
P_6

$$A_1^T \boldsymbol{x} \geq \boldsymbol{b}_1$$
$$A_5^T \boldsymbol{x} \leq \boldsymbol{b}_5$$

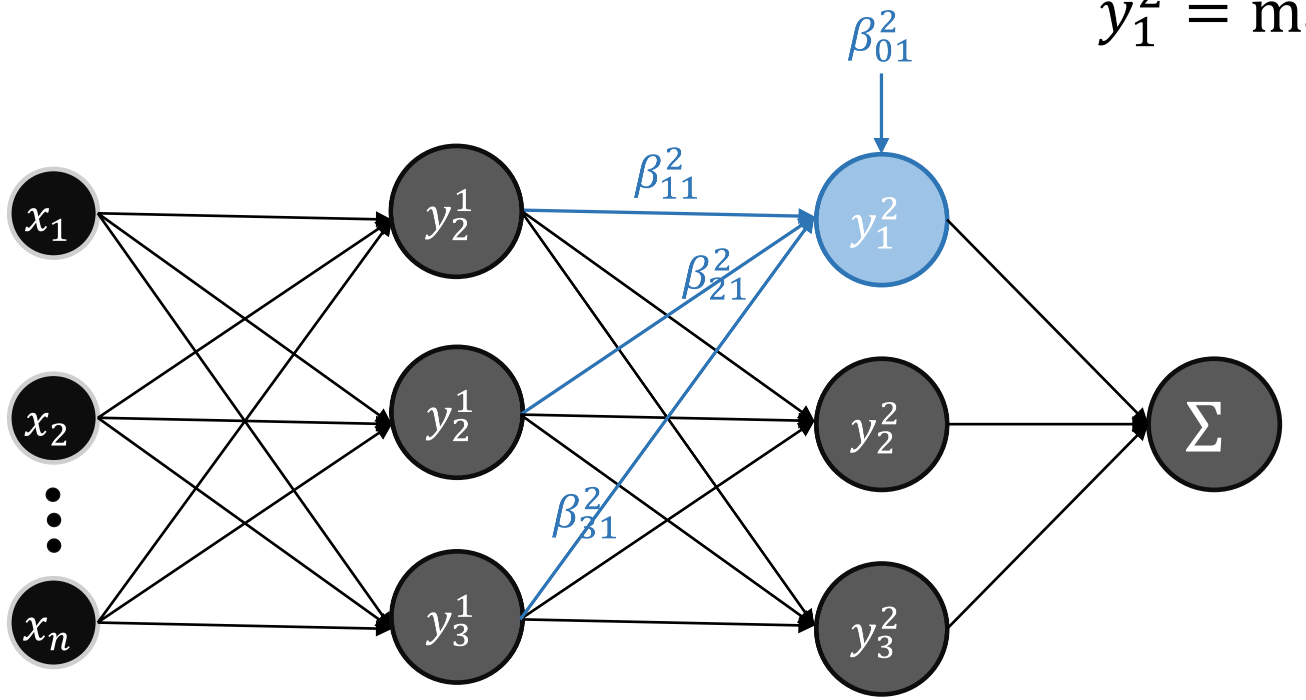
Embedding Neural Networks

ReLU activation function

$$y_1^2 = \max\left\{0, \beta_{01}^2 + \boldsymbol{\beta}_1^{2^T} \boldsymbol{y}^1\right\}$$



Embedding Neural Networks



ReLU activation function

$$y_1^2 = \max\left\{0, \beta_{01}^2 + \boldsymbol{\beta}_1^{2^T} \boldsymbol{y}^1\right\}$$

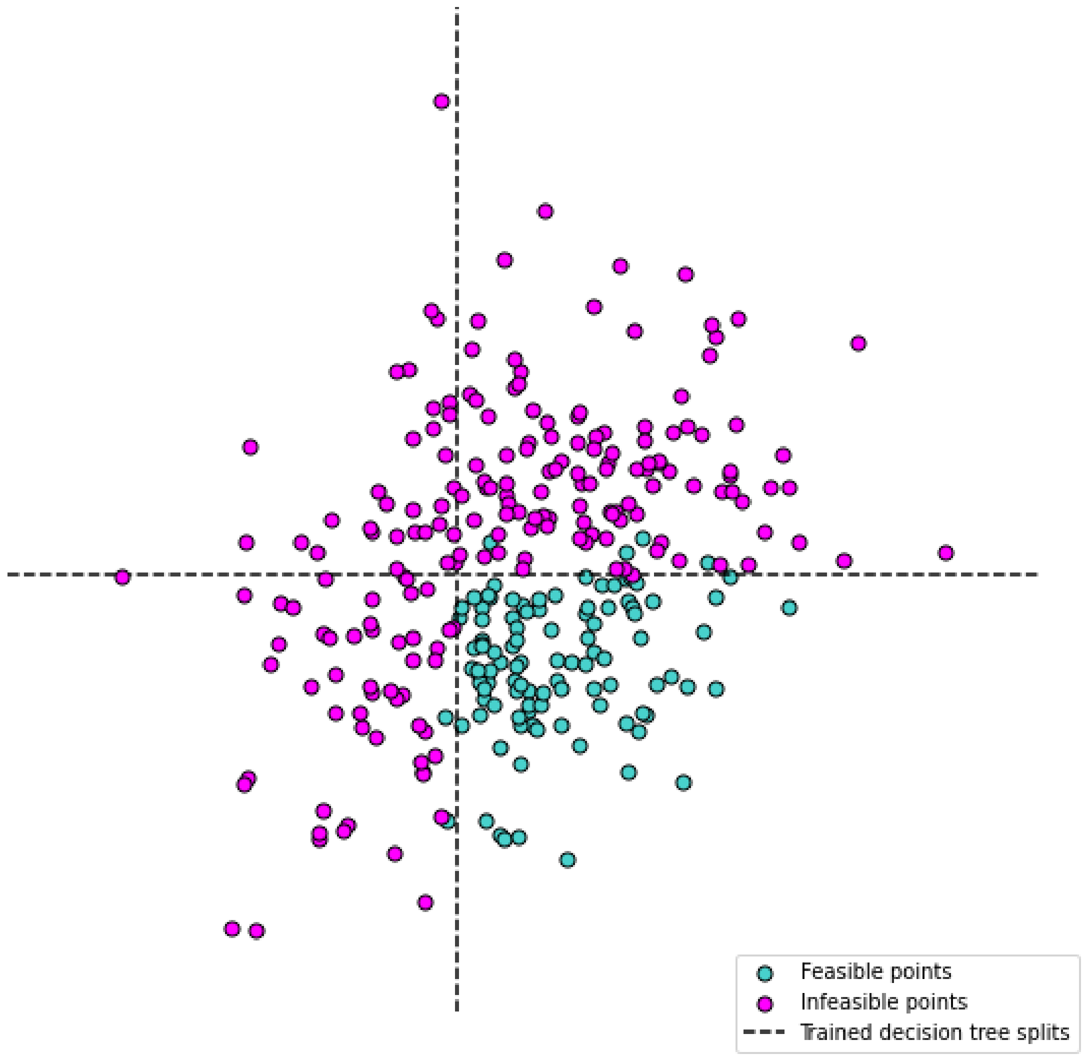
$y = \max\{0, x\}$ can also be written as:

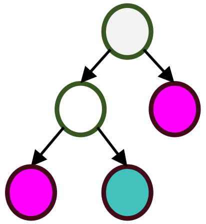
$$\begin{cases} y \geq x, \\ y \leq x - M_L(1 - z), \\ y \leq M_U z, \\ y \geq 0, \\ z \in \{0, 1\}, \end{cases}$$

where $M_L < 0$ is a lower bound on all possible values of x, and $M_U > 0$ is an upper bound.

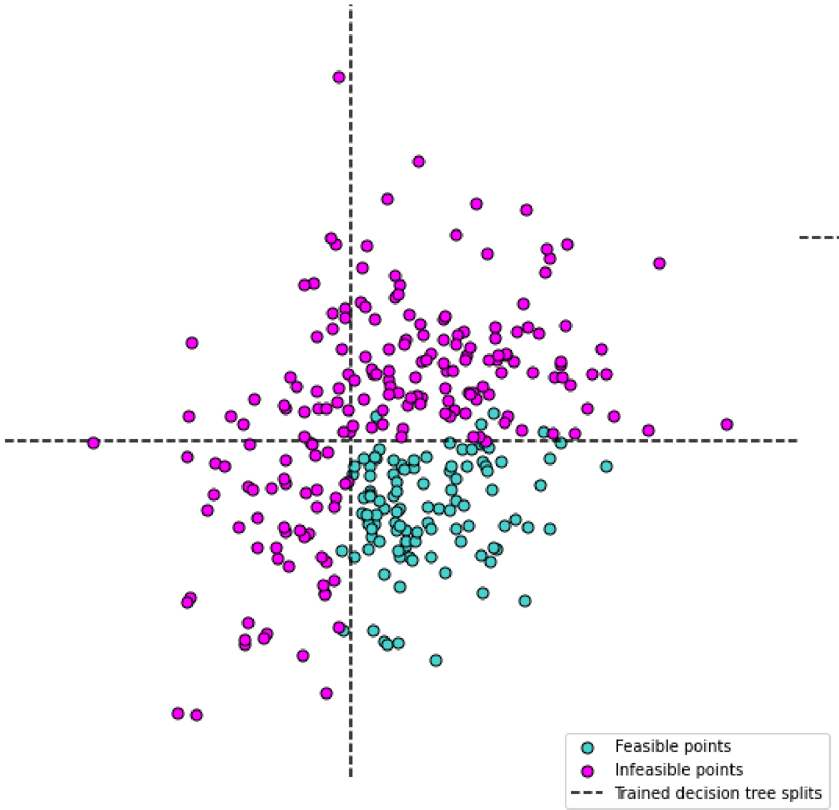
Trust region constraints

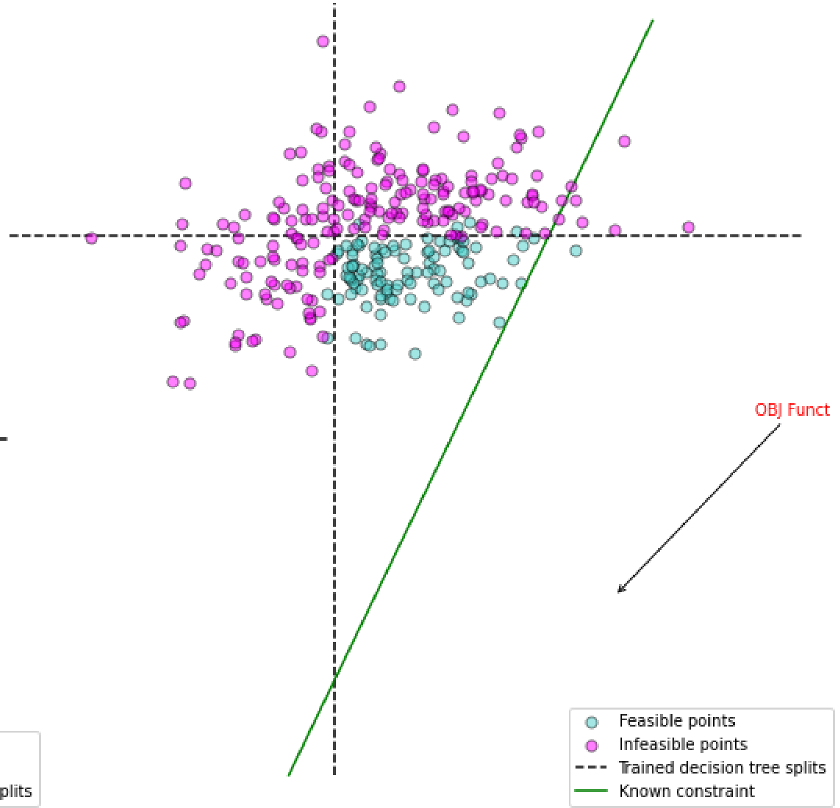
UNIVERSITY OF AMSTERDAM IBM

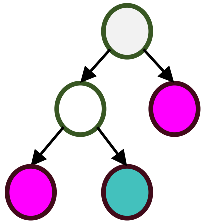




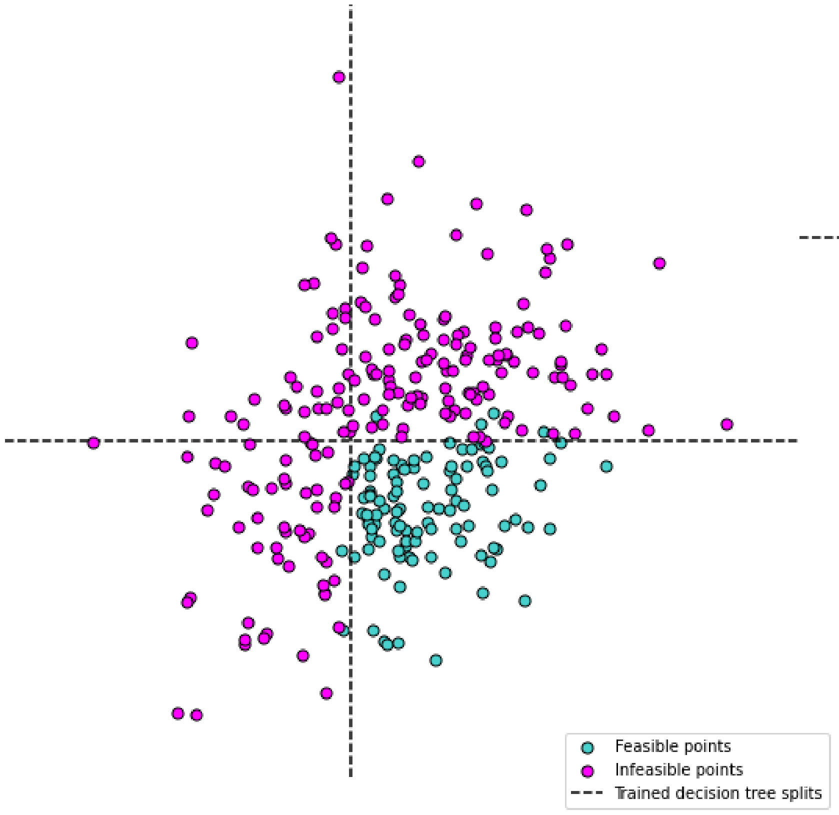
Trust region constraints

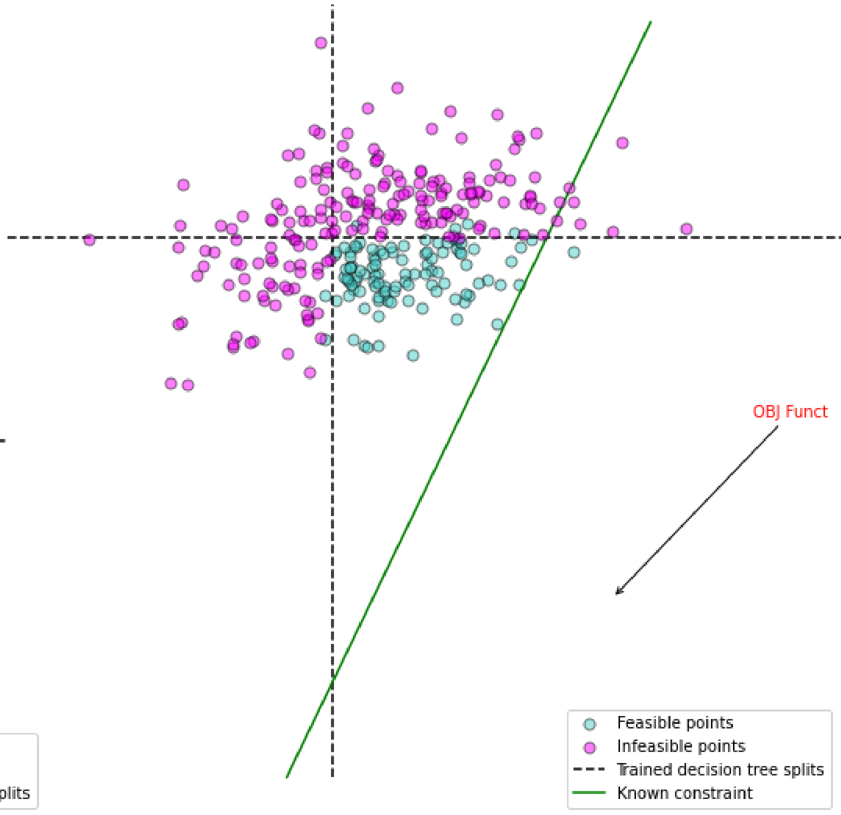


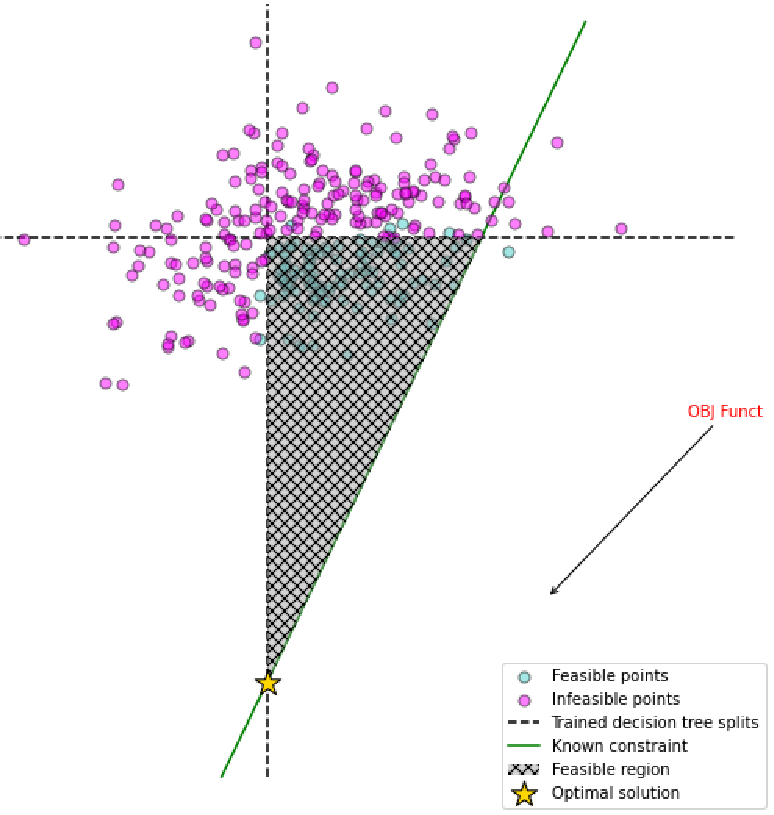


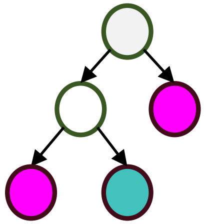


Trust region constraints

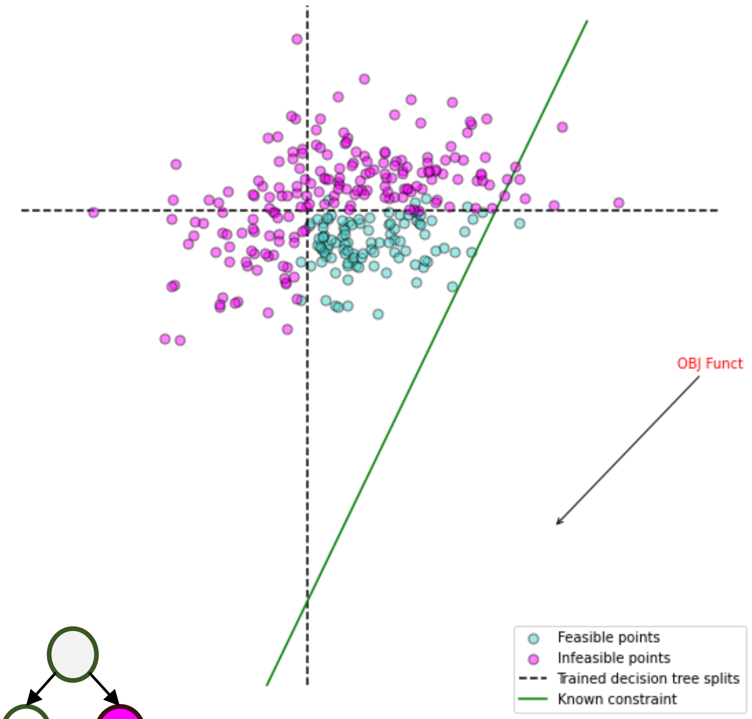


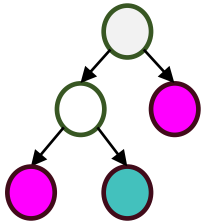






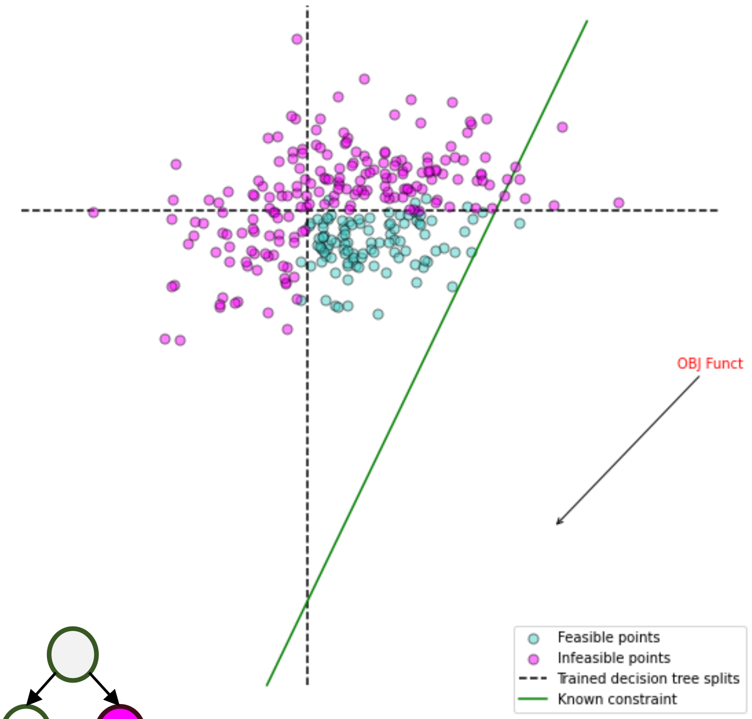
Trust region constraints

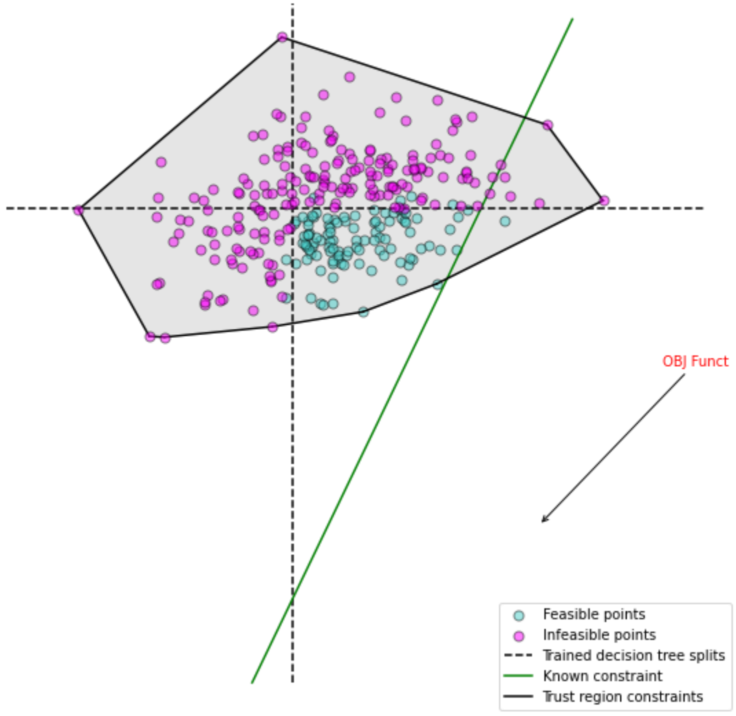


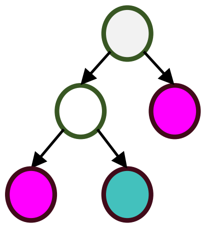


Trust region constraints

$$CH(\boldsymbol{x}) = \left\{ \boldsymbol{x} \,|\, \boldsymbol{x} = \sum_{i}^{N} \lambda_i \overline{\boldsymbol{x}}_i \,, \sum_{i}^{N} \lambda_i = 1 \,, \lambda_i \geq 0, i = 1, \dots, N \right\}$$

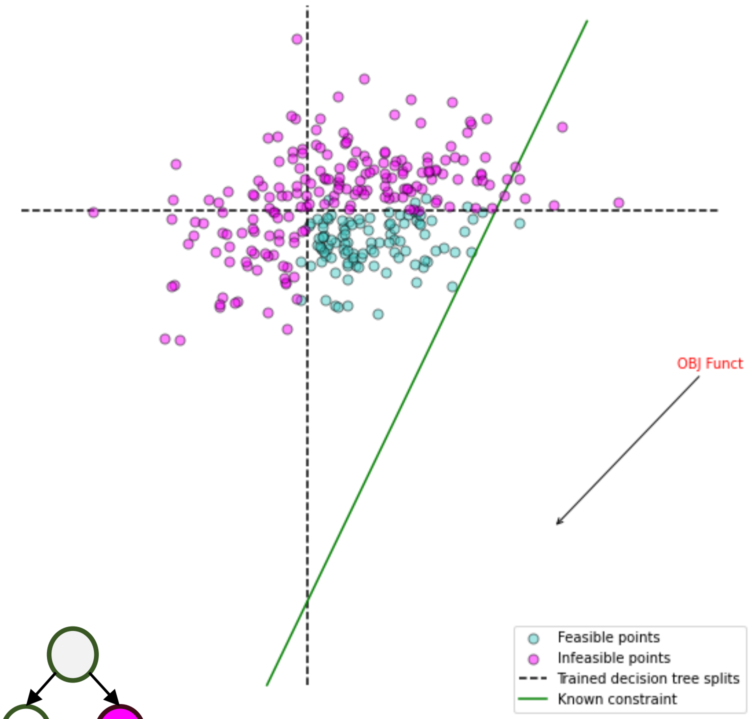


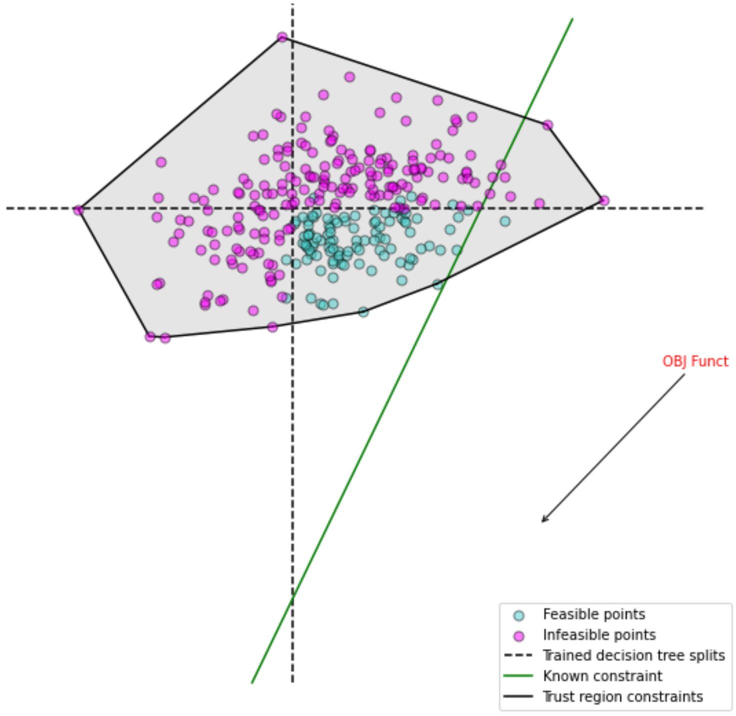


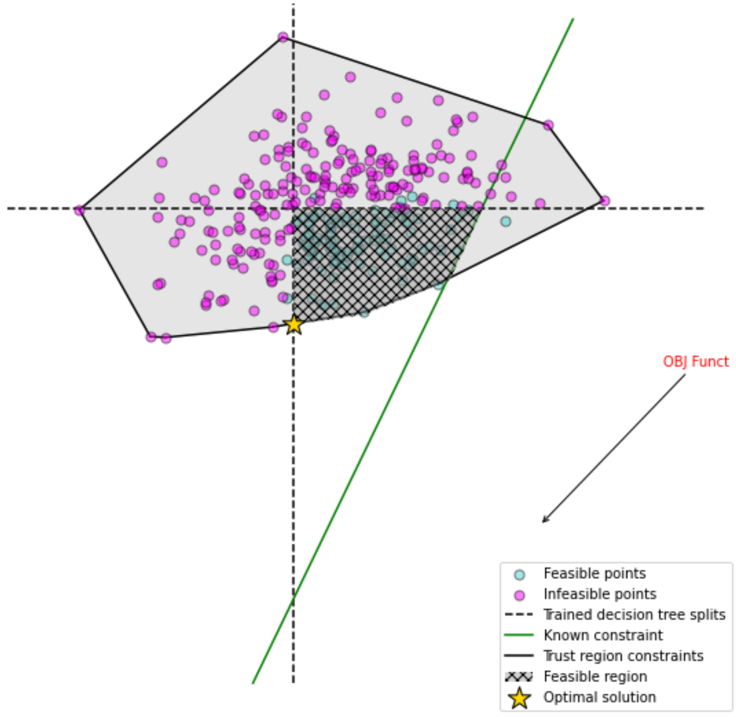


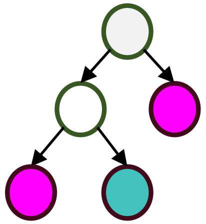
Trust region constraints

$$CH(\boldsymbol{x}) = \left\{ \boldsymbol{x} | \boldsymbol{x} = \sum_{i}^{N} \lambda_i \bar{\boldsymbol{x}}_i \, , \sum_{i}^{N} \lambda_i = 1 \, , \lambda_i \geq 0, i = 1, \ldots, N \right\}$$









OptiCL

A Python Package for
Optimization with Constraint Learning

https://github.com/hwiberg/OptiCL

UNIVERSITY OF AMSTERDAM IBM

OptiCL

A Python Package for

Optimization with **C**onstraint **L**earning

https://github.com/hwiberg/OptiCL

scikit learn

PYOMO

Hands-on tutorial on the

diet problem

WFP

Thank you!

Q&A

CONTACT ME

d.maragno@uva.nl

github.com/donato-maragno

@MaragnoDonato

linkedin.com/in/donato-maragno