

Optimization meets Machine Learning

Marco Lübbecke
Lehrstuhl für Operations Research
RWTH Aachen University



@mluebbecke

OR 2017 · Berlin · September 6, 2017

Optimization is Everywhere



Machine Learning is Everywhere



Machine Learning is Everywhere



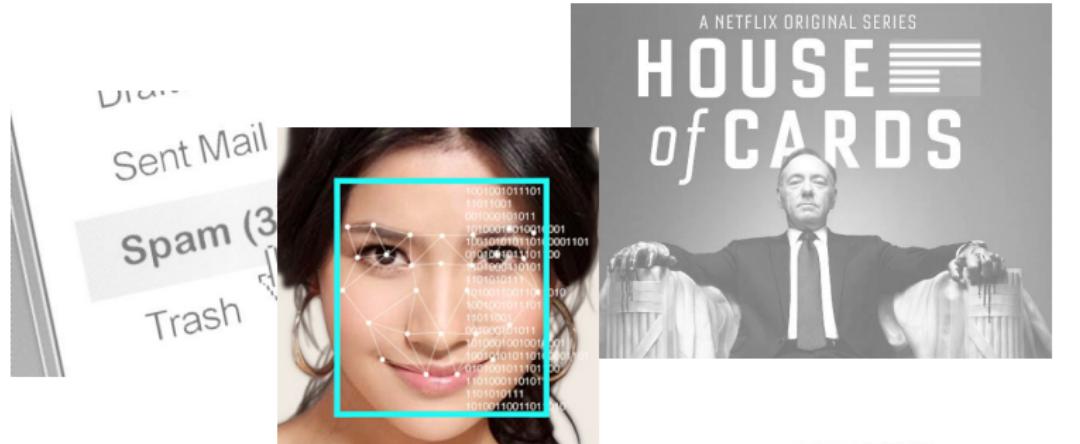
Machine Learning is Everywhere



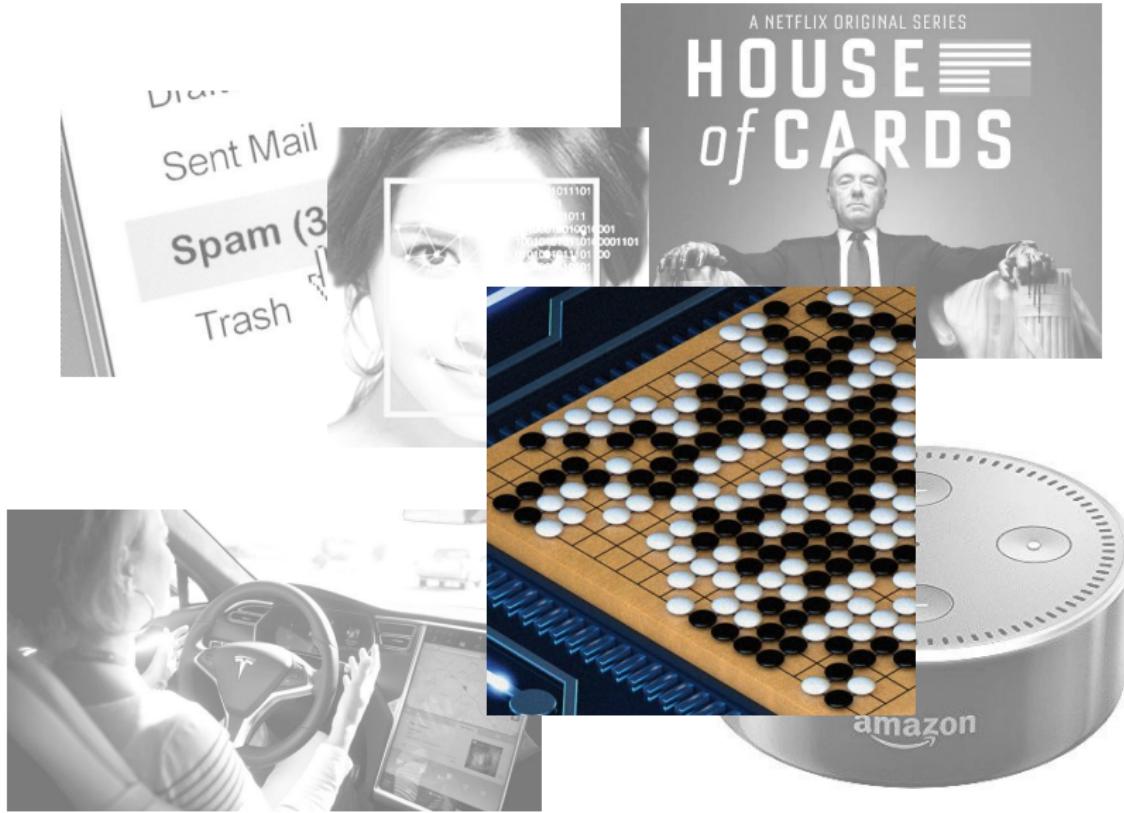
Machine Learning is Everywhere



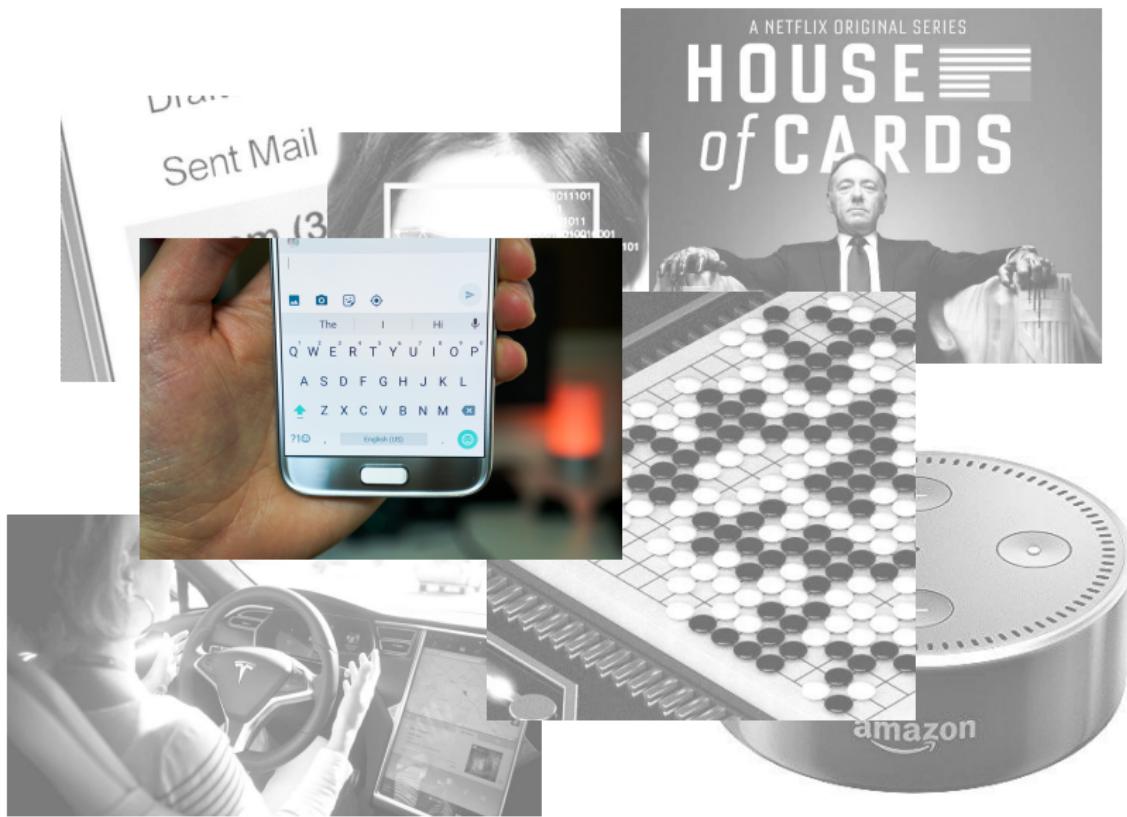
Machine Learning is Everywhere



Machine Learning is Everywhere



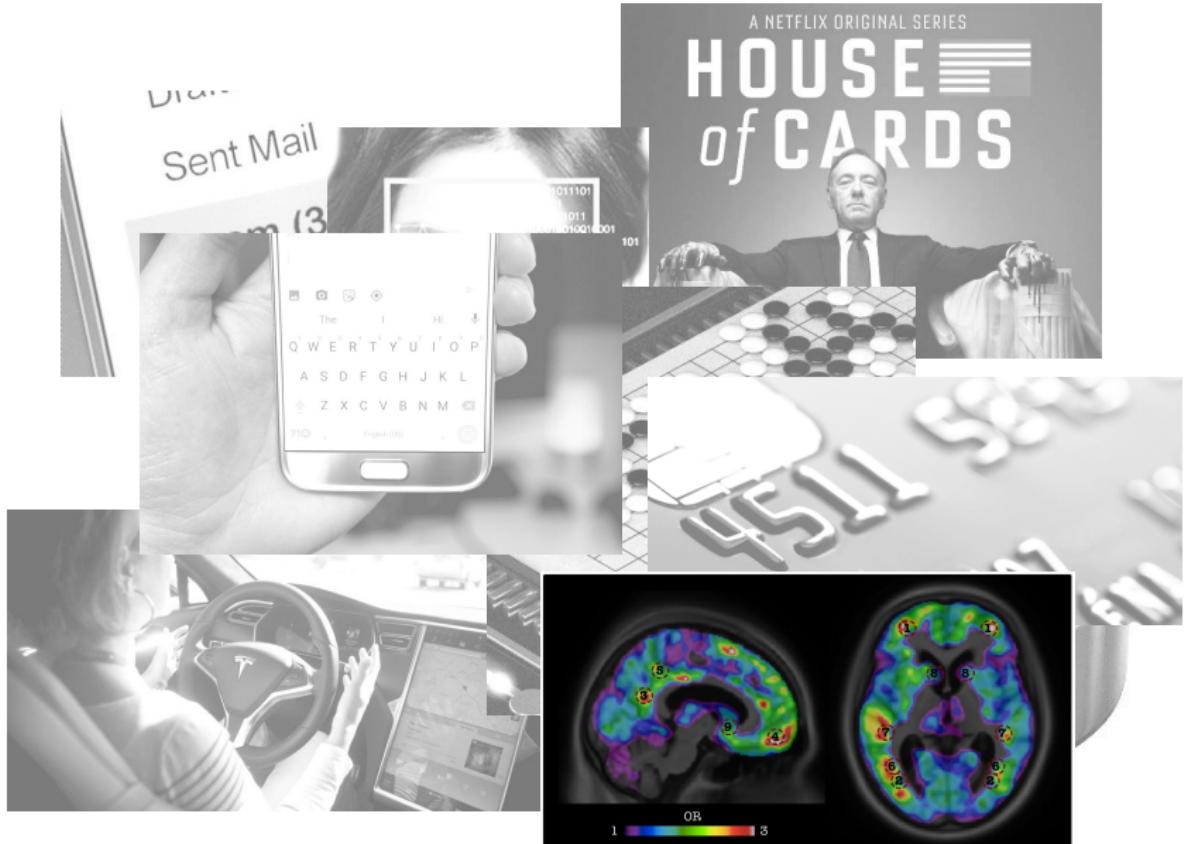
Machine Learning is Everywhere



Machine Learning is Everywhere



Machine Learning is Everywhere



Literally *Everyone* speaks about Machine Learning

The screenshot shows the homepage of DER TAGESSPIEGEL. At the top, there are social media icons for Facebook, Twitter, and Instagram, along with 'LOGIN' and 'REGISTRIEREN' buttons. Below this is the newspaper's logo 'DER TAGESSPIEGEL' with a small globe icon. A navigation bar below the logo includes links for LITIK, BERLIN, WIRTSCHAFT, SPORT, KULTUR, WELT, MEINUNG, MEDIEN, WISSEN, QUEER, VERBRAUCHER, and a search bar with a magnifying glass icon. The main content area shows a breadcrumb trail: Home > Digital Science Match > Digital Humanities | Digital Media Services > Prof. Dr. Klaus-Robert Müller: Machine Learning and Applications. The article is dated 02.11.2015 13:23 Uhr and is written by Prof. Dr. Klaus-Robert Müller. The title of the article is 'Machine Learning and Applications'. Below the title, a subtitle reads: 'Buchstabieren und Spielen nur mit den Gedanken - mit dem Berliner Brain Computer Interface lernt eine Maschine Hirnsignale zu dekodieren.' A large thumbnail image of a man in a suit, identified as the author, is displayed.

Literally *Everyone* speaks about Machine Learning

The screenshot shows a news article from Zeit Online. At the top, there's a navigation bar with links for ABO, SHOP, AKADEMIE, JOBS, MEHR, LOGIN, and REGISTRIEREN. Below that is the Zeit Online logo and a search bar. The main headline is "Künstlich, intelligent und frei" by Eike Kühl. The text discusses how companies like IBM, Facebook, and Google invest in artificial intelligence research. A small image at the bottom shows a group of people and animals.

DER TAGESSPIEGEL

ABO SHOP AKADEMIE JOBS MEHR •

E-PAPER AUDIO APPS ARCHIV ANMELDEN

ZEIT ONLINE

Suche

Politik Gesellschaft Wirtschaft Kultur • Wissen Digital Campus • Karriere Entdecken Sport ZEITmagazin mehr • #D17

Prof. Dr. K.

Mach

Buchstaben
eine Mas

Maschinelles Lernen

Künstlich, intelligent und frei

Firmen wie IBM, Facebook und Google investieren Millionen in künstliche Intelligenz. Teile ihrer Forschung stellen sie für alle frei zur Verfügung – aus gutem Grund.

Von **Eike Kühl**

14. Dezember 2015, 19:28 Uhr / 19 Kommentare



@mluebbecke

Optimization meets Machine Learning

4/30

Literally *Everyone* speaks about Machine Learning

The screenshot shows a news article from the website of t3n magazine. The main headline is "So verändert Machine Learning die Wirtschaft". The article has 155 shares and includes social sharing buttons for Facebook, Twitter, and LinkedIn. Below the headline is a large graphic of a lightbulb with a brain inside, symbolizing ideas or innovation. The article is categorized under "Politik & Gesellschaft" and "Digital". The URL of the article is t3n.de/magazin/so-veraendert-machine-learning-die-wirtschaft/.

DER TAGESSPIEGEL

ABO SHOP AKADEMIE JOBS MEHR *

ZEIT

LITIK BERLIN

Politik Gesell

Prof. Dr. K.

Mach

Buchstabe einer Mas

155 SHARES / TEILEN TWITTER TEILEN IN TEILEN MAILEN

t3n 46

So verändert Machine Learning die Wirtschaft

Autor: Boris Hänsler

erschienen in t3n 46

RIP REST IN PEACE 12/2016 - 02/2017 JETZT KAUFEN

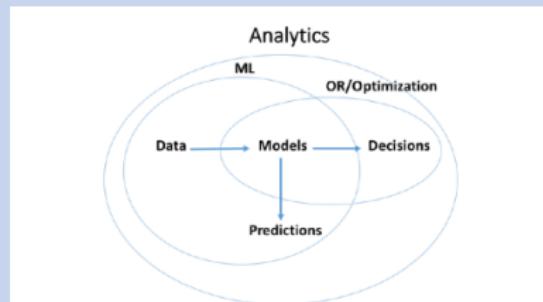
Die besten News per E-Mail! Sichere dir deinen Wissenvorsprung!

Aktuelle News Exklusive Goodies Werbezeit kündbar

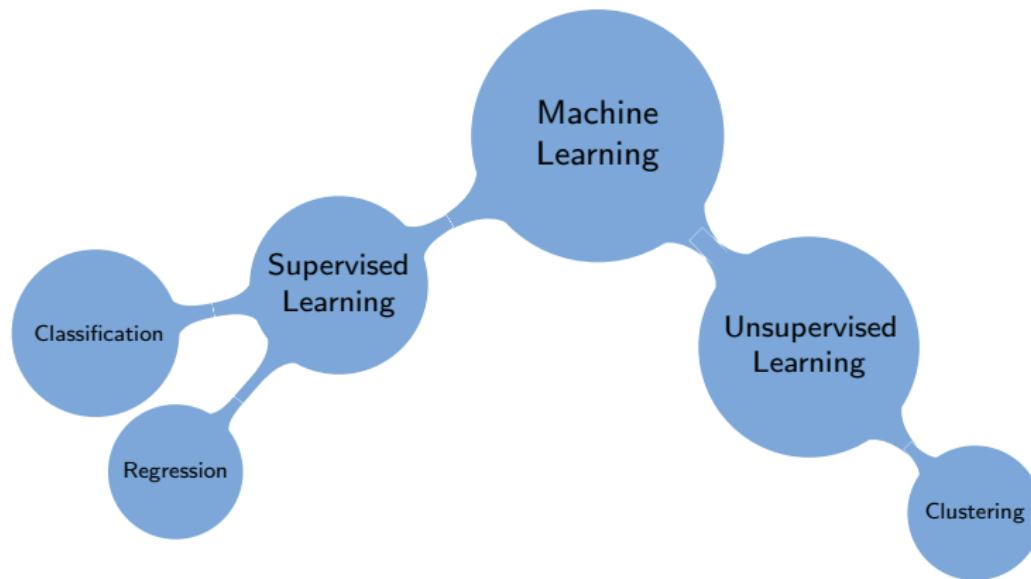
New INFORMS Journal on Optimization

“ One of the largest opportunities of the field of optimization is to embrace data in a protagonist role and combine it with machine learning.
— Dimitris Bertsimas, Editorial Statement ”

“ My vision of the future for [...] optimization



What *is* Machine Learning?



source: www.mathworks.com

Supervised Learning: Classification

► data \mathcal{X}



Supervised Learning: Classification

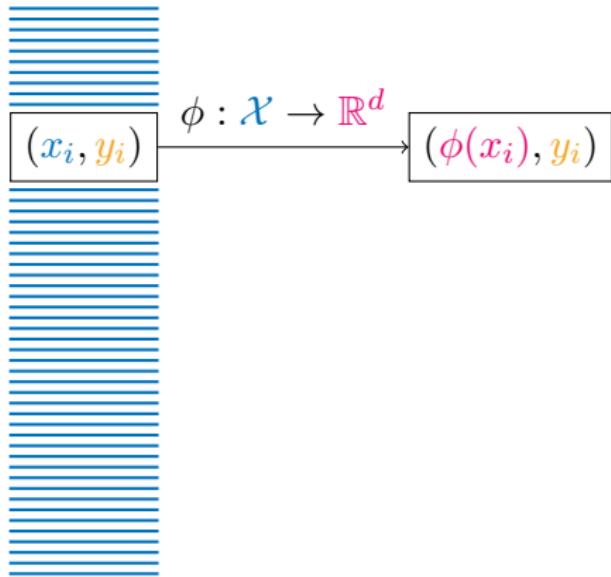
► data \mathcal{X}

labels \mathcal{Y}



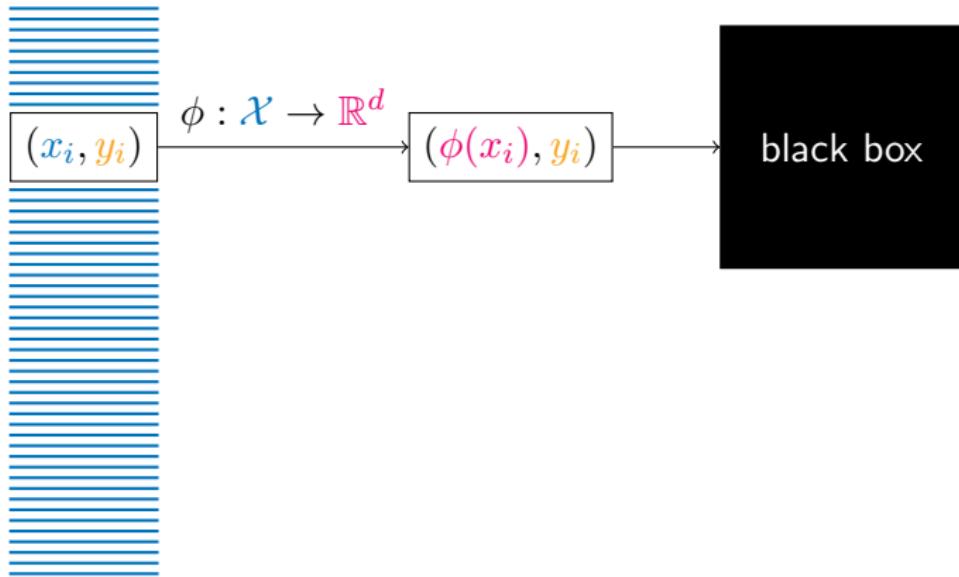
Supervised Learning: Classification

- ▶ data \mathcal{X} , d features, labels \mathcal{Y}



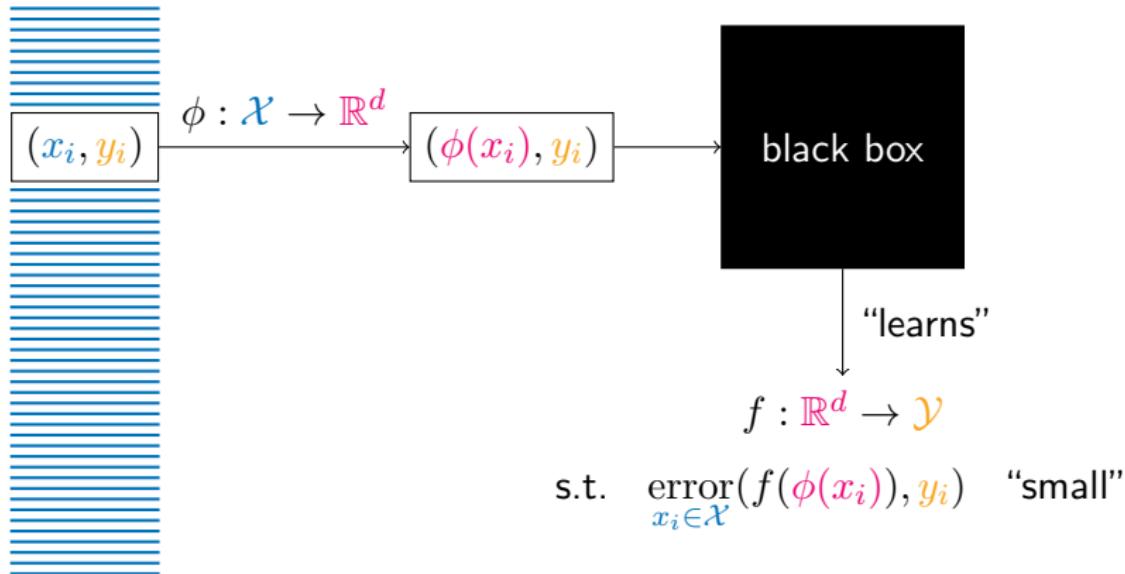
Supervised Learning: Classification

- ▶ data \mathcal{X} , d features, labels \mathcal{Y}



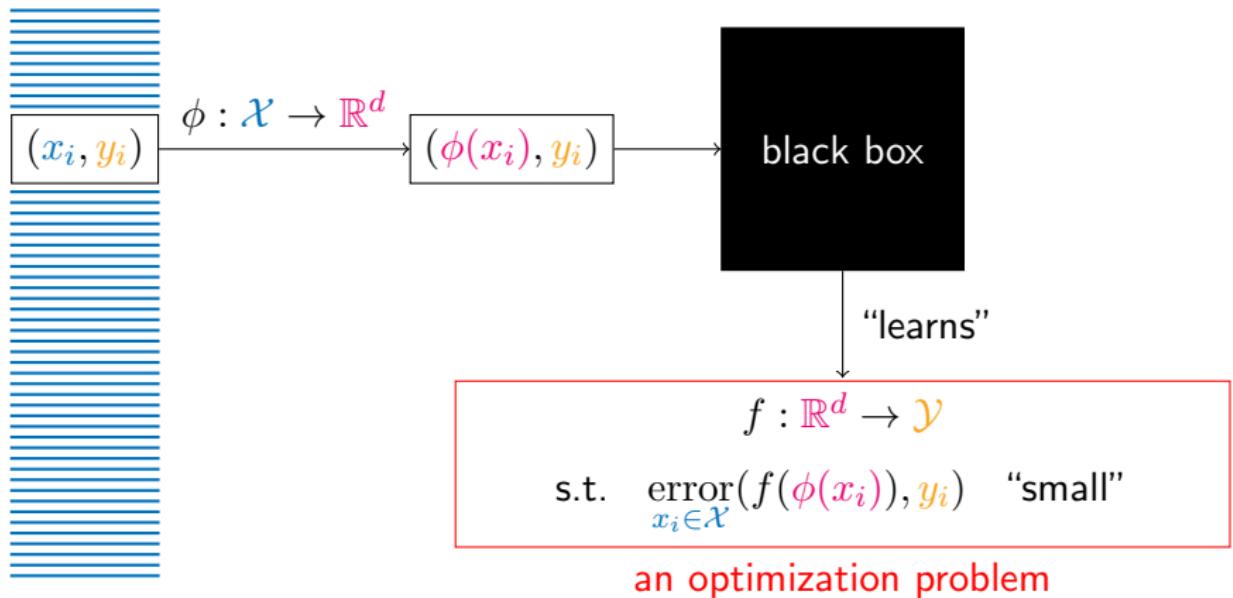
Supervised Learning: Classification

- ▶ data \mathcal{X} , d features, labels \mathcal{Y}



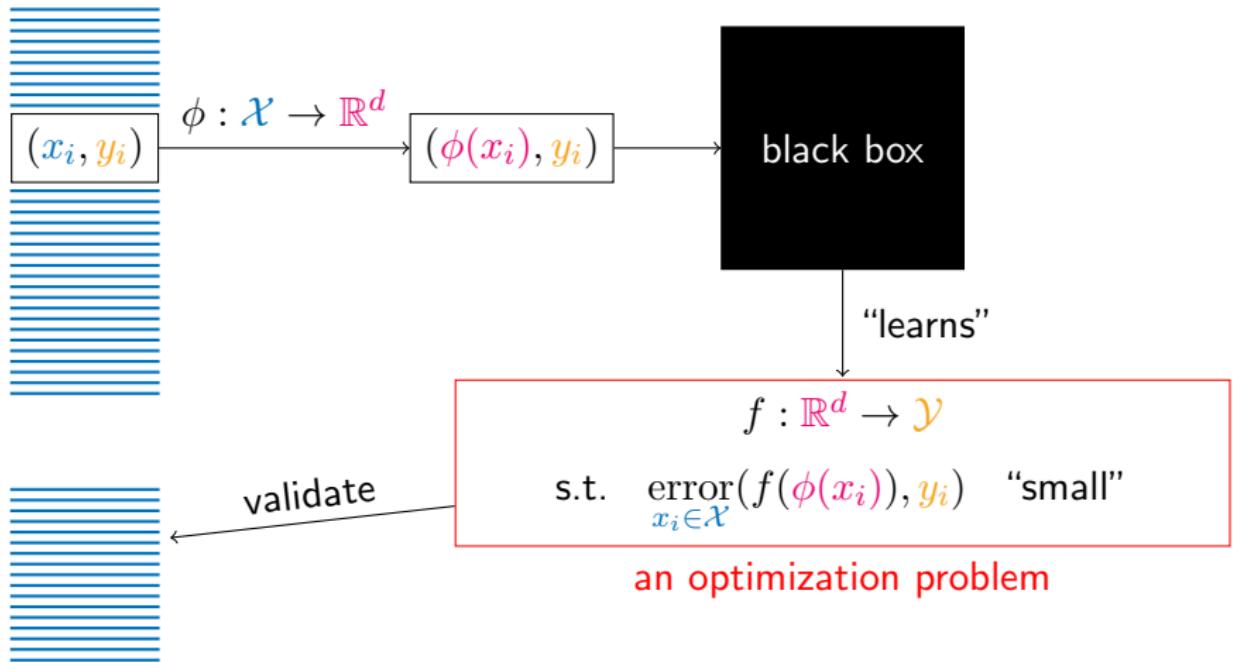
Supervised Learning: Classification

- ▶ data \mathcal{X} , d features, labels \mathcal{Y}



Supervised Learning: Classification

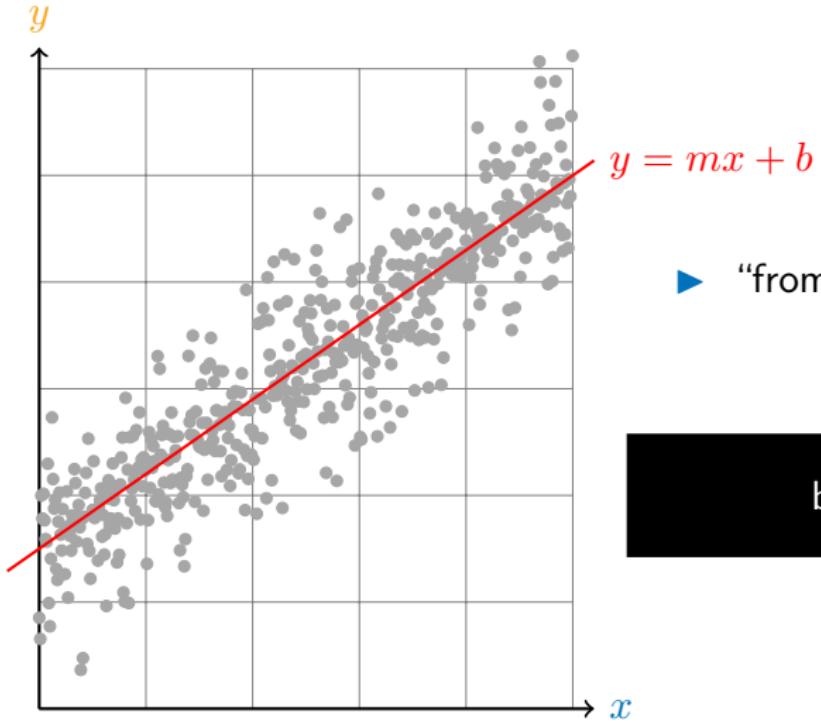
- ▶ data \mathcal{X} , d features, labels \mathcal{Y}



Binary Classification: Dog or Muffin? Owl or Apple?



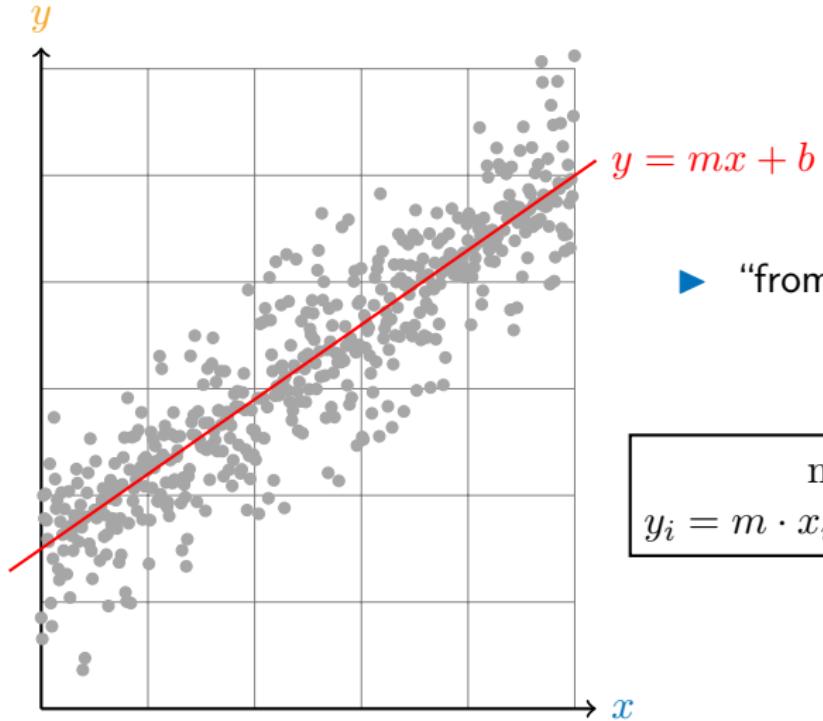
Supervised Learning: Regression



► “from $\{0, 1\}$ to $[0, 1]$ ”

black box

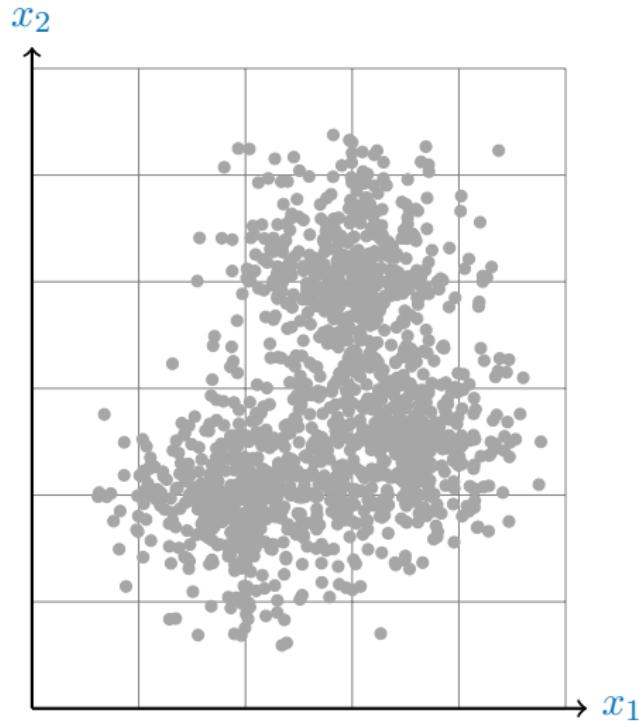
Supervised Learning: Regression



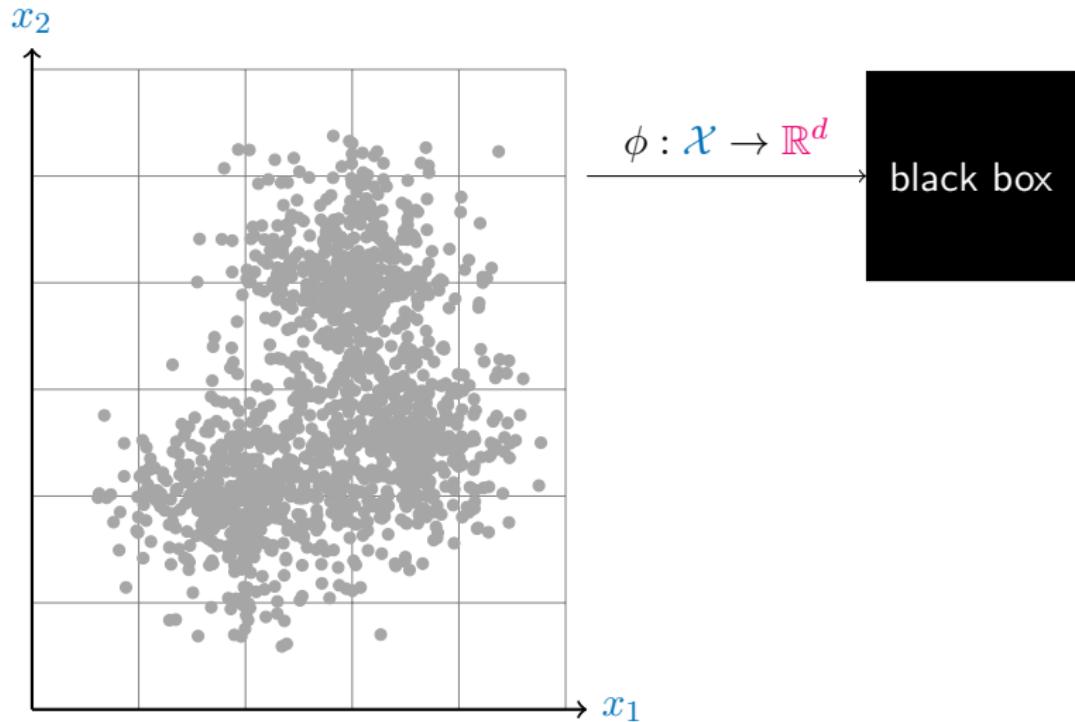
► “from $\{0, 1\}$ to $[0, 1]$ ”

$$\min \sum_i \varepsilon_i^2$$
$$y_i = m \cdot x_i + b + \varepsilon_i \quad x_i \in \mathcal{X}$$

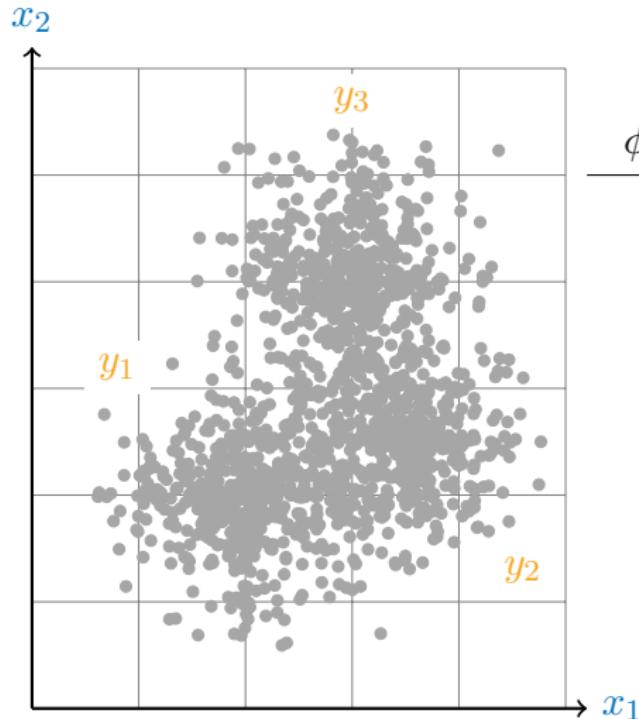
Unsupervised Learning: Clustering



Unsupervised Learning: Clustering



Unsupervised Learning: Clustering



$$\phi : \mathcal{X} \rightarrow \mathbb{R}^d$$



black box

“learns”

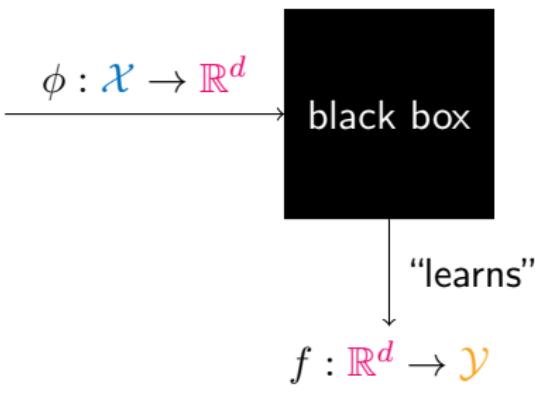
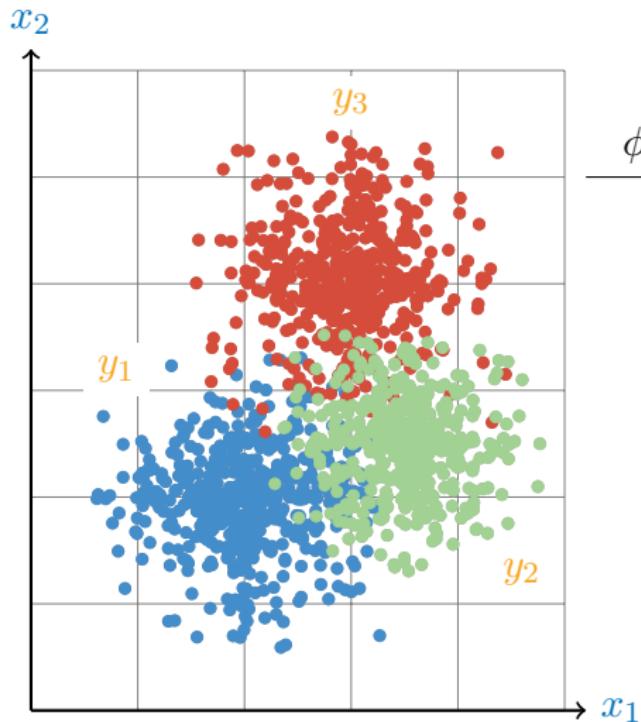
$$f : \mathbb{R}^d \rightarrow \mathcal{Y}$$

s.t. all

$$x \in \mathcal{X} : f(\phi(x)) = y$$

are “similar”

Unsupervised Learning: Clustering



s.t. all
 $x \in \mathcal{X} : f(\phi(x)) = y$
 are “similar”

Optimization naturally appears in ML

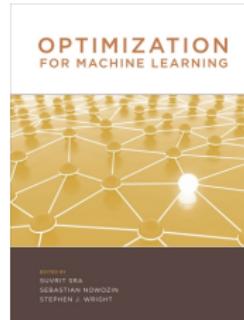


Optimization lies at the heart of ML. Most ML problems reduce to optimization problems.

— Bennett, Parrado-Hernández (2006)

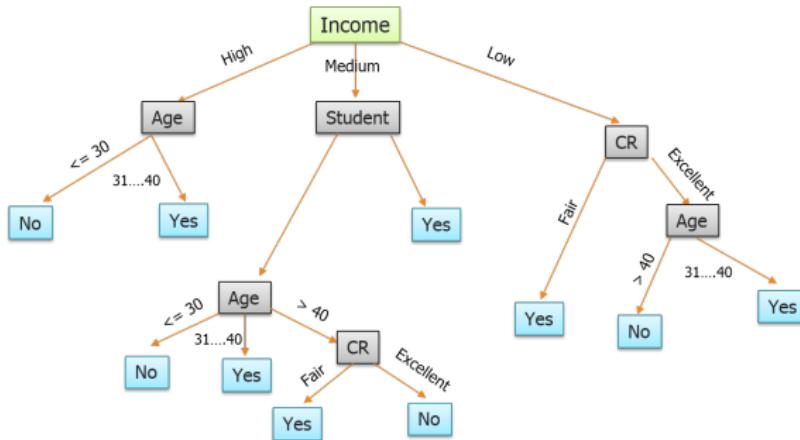


- ▶ minimize e.g., prediction error
- ▶ continuous, convex optimization
- ▶ discrete, integer optimization



Example: A MIP in a black box : Classification Trees

- optimal classification trees Bertsimas & Dunn (2017)



source: www.edureka.co/blog/decision-trees

- use few nodes, shallow depth → formulated as MIP
- improves accuracy over classical CART method by 0.5–2%

Many Opportunities for Discrete Optimization in ML

- ▶ within **black boxes** to capture combinatorial explosion
- see also Andrea's plenary on Friday
- ▶ feature selection
- ▶ outlier detection
- ▶ parameter tuning
- ▶ ...

How about the converse Direction?

- ▶ “emulating the expert” Bärmann, Pokutta & Schneider (2017)
- ▶ observe a decision maker $\max c_{\text{true}}^T x : x \in X(p)$
- ▶ learn their objective function given $(p_t, x_t^*)_{t=1,\dots,T}$
- ▶ *online learning*

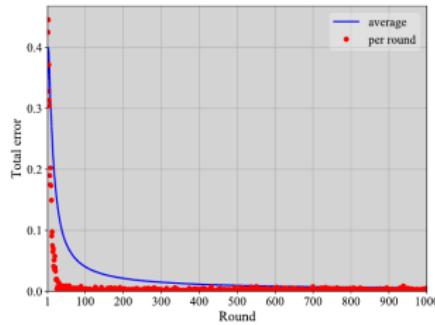


Figure 1. Linear Knapsack problem with $n = 100$ items over

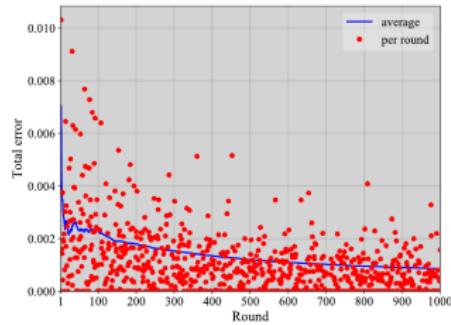


Figure 2. Resource-constrained shortest path problem on a grid

ML may help improving Optimization Algorithms

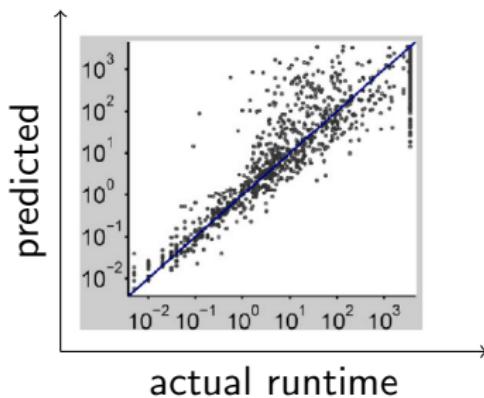
- ▶ e.g., branching in B&B
- ▶ *full strong branching* gives locally perfect information
- ▶ predict the **strong branching score** of a variable
- ▶ **features** describe **state of a variable**
- ▶ supervised learning: regression
- promising proof-of-concept

Marcos Alvarez, Louveaux & Wehenkel (2017)

- ▶ survey on ML in branching/searching [Lodi & Zarpellon \(2017\)](#)

A Progress Bar for Branch-and-Bound?

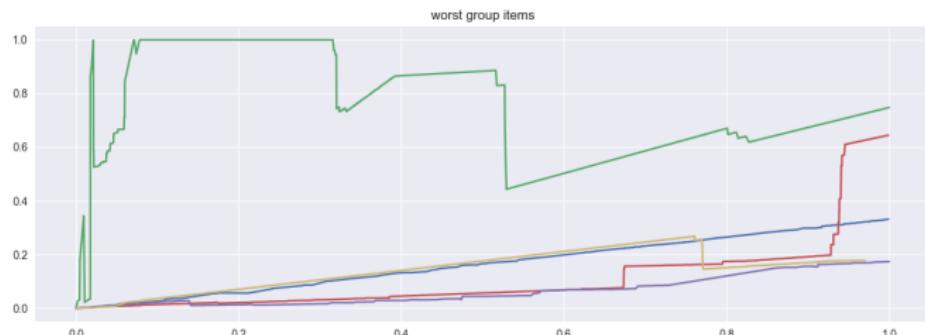
- ▶ predict the runtime of branch-and-bound algorithms
[Hutter, Xu, Hoos & Leyton-Brown \(2014\)](#)
- ▶ CPLEX 12.1 on 1510 publicly available MIPs



A Progress Bar for Branch-and-Bound?

- ▶ very preliminary experiments with gurobi 7.5
- ▶ predict elapsed runtime percentage

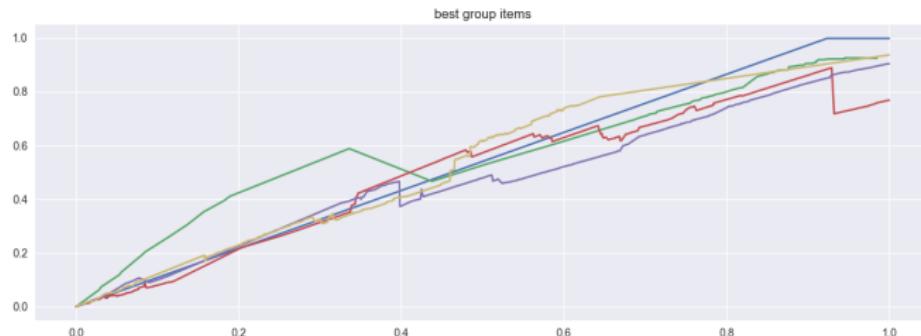
Kruber, L, Obeloeer genannt Bregenhorst (2017)



A Progress Bar for Branch-and-Bound?

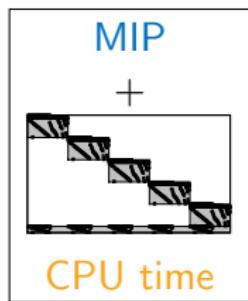
- ▶ very preliminary experiments with gurobi 7.5
- ▶ predict elapsed runtime percentage

Kruber, L, Obeloeer genannt Bregenhorst (2017)



Learning when to solve a MIP by Branch-and-Price

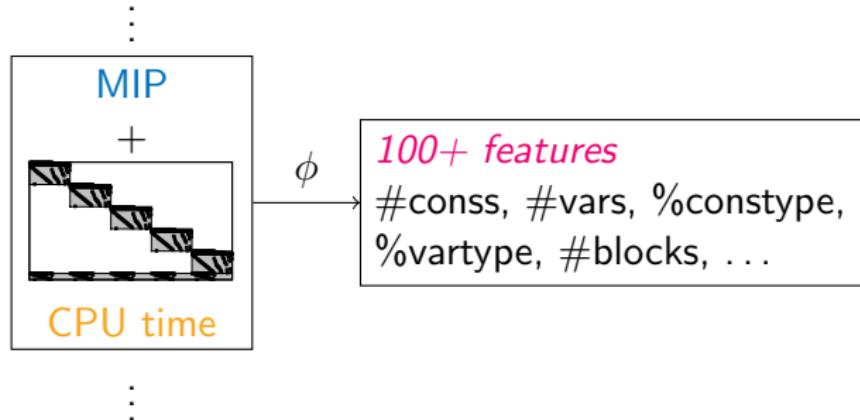
- ▶ our MIP solver GCG detects many potential DW reformulations



Kruber, L, Parmentier (2017)

Learning when to solve a MIP by Branch-and-Price

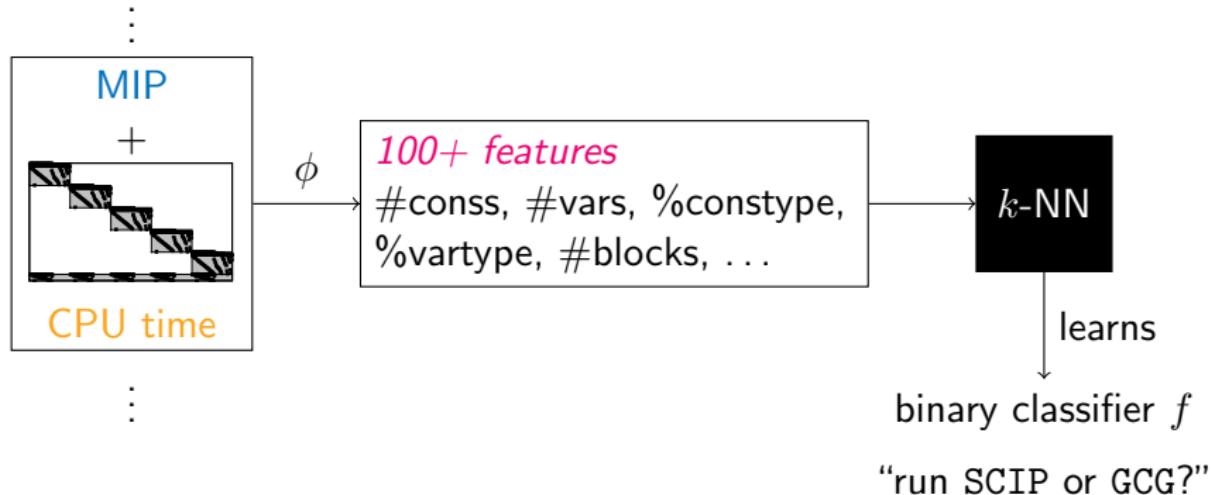
- ▶ our MIP solver GCG detects many potential DW reformulations



Kruber, L, Parmentier (2017)

Learning when to solve a MIP by Branch-and-Price

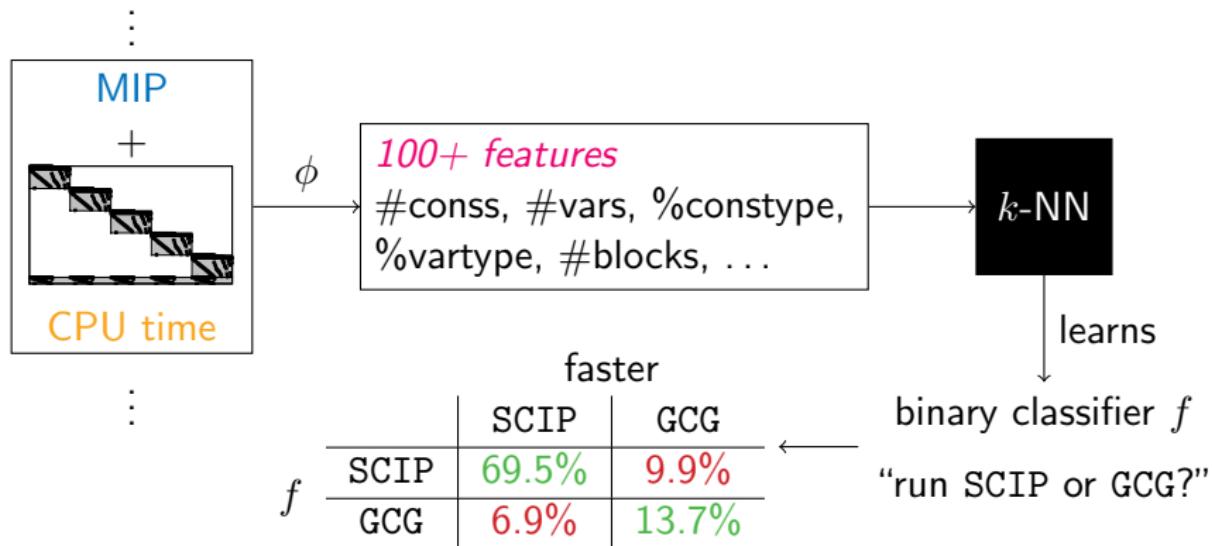
- ▶ our MIP solver GCG detects many potential DW reformulations



Kruber, L, Parmentier (2017)

Learning when to solve a MIP by Branch-and-Price

- ▶ our MIP solver GCG detects many potential DW reformulations



Kruber, L, Parmentier (2017)

What ML Answers can we (Optimizers) expect?

- ▶ we get statistical answers → not what we are used to see
 - ▶ we have domain/expert knowledge: e.g., pseudo-costs
 - ▶ ML may give a better predictor, but no *explanation*
 - ▶ some info can be extracted from most influential **features**
- **interpretability** is a huge theoretical and practical topic



**INTERNATIONAL CONFERENCE
ON OPERATIONS RESEARCH**

Decision Analytics for the Digital Economy

FREIE UNIVERSITÄT BERLIN

SEPTEMBER 06 – 08

www.or2017.de



Decision Making: Machine Learning



Machine Learning and Artificial Intelligence delivers the most value when you need to make lots of similar decisions quickly.

— Ingo Mierswa, Rapidminer



- ▶ **simple decisions:** e.g., auto correct current word
- ▶ **solution:** often a single score → greedy
- ▶ **keep/learn habits:** extrapolate from the past (!)

Typical Example: Predictive Maintenance



source: blog.capterra.com/should-you-invest-in-a-predictive-maintenance-strategy/

Exploit all Options: Prescriptive Maintenance



source: www.siemens.com/press/pool/de/pressebilder/photonews/pn200826/300dpi/pn200826-12_300dpi.jpg

Decision Making: Optimization

“ “ How often can the result of an optimization model be captured in a single variable?
— Ed Rothberg, Gurobi “ ”

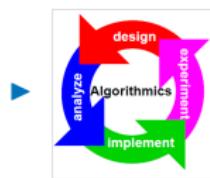
- ▶ **solution:** not only the objective value!
- ▶ **complex decisions/plans:** e.g., timetables, crew schedules, ...
- ▶ **global scope:** models all (reasonable) interdependencies

"Current Standard:" Predictive then Prescriptive Analytics

ML harnesses the bigness of data (the past and present);
Optimization captures the bigness of options (the future).

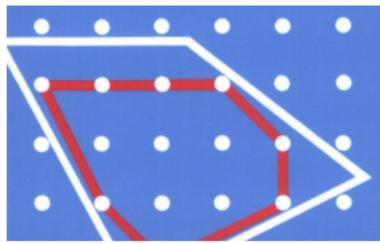
Learning (about) optimal Solutions

- ▶ in recurring complex decision situations
- ▶ learn how good (partial) solutions look like
- ▶ this may help finding good solutions faster
- ▶ learn spatio-temporal patterns to generate effective schedules
[Le, Liu, Lau \(2016\)](#)



Learning (about) Optimization Models

- ▶ ML can make sense of data
- ▶ optimization models are also “data”
- (how) can ML help us make sense of optimization models?



- ▶ can ML learn good modeling?

Vision: Learning (about) Optimization Problems

- ▶ learn the *semantics* of a MIP model (“the problem”)

$$\begin{aligned} & \min \sum_{j=1}^n y_j \\ \text{s.t. } & \sum_{j=1}^n x_{ij} = 1 \quad i = 1, \dots, n \\ & \sum_{i=1}^n a_{ij} x_{ij} \leq b \quad j = 1, \dots, n \\ & x_{ij} \leq y_j \quad i, j = 1, \dots, n \\ & x_{ij}, y_j \in \{0, 1\} \quad i, j = 1, \dots, n \end{aligned}$$



$$\begin{aligned} & \min \sum_{p \in P} \nu_p \\ \text{s.t. } & \sum_{p \in P: i \in p} \nu_p = 1 \quad i = 1, \dots, n \\ & \nu_p \in \{0, 1\} \quad p \in P \end{aligned}$$

⇒ e.g., help the modeler find a better formulation

The AI Umbrella?

- ▶ OR vs. analytics discussion ✓
- ▶ OR vs. AI discussion ?



Marco Lübecke
@mluebbecke

#artificialintelligence: a machine mimics "cognitive" functions like learning and problem solving. would you consider #orms as #AI?

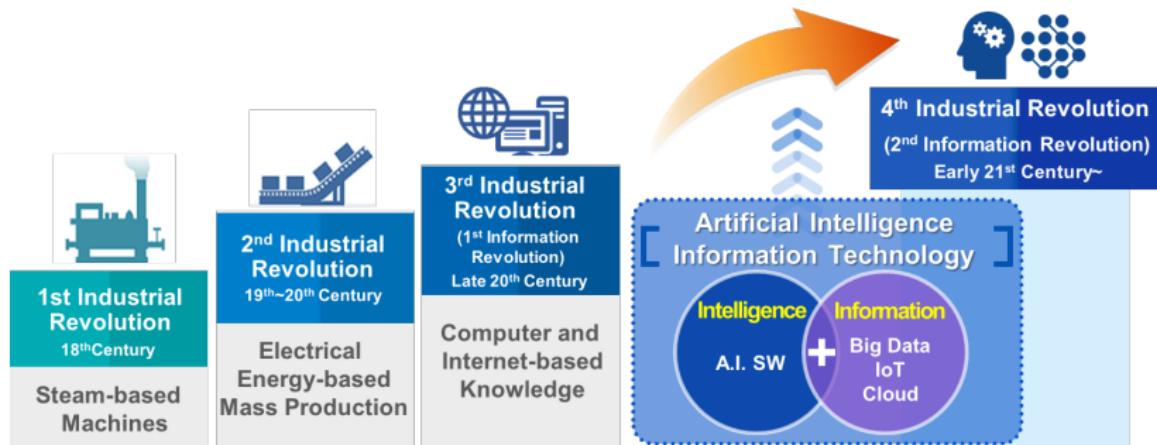
Original (Englisch) übersetzen



35 Votes • Endergebnisse

12:43 nachm. · 28 Aug. 17

Why is this Relevant?



source: blogs.worldbank.org/category/tags/artificial-intelligence

- ▶ if the fourth industrial revolution is about AI,
OR should be part of it

Optimization met Machine Learning

Marco Lübbecke
Lehrstuhl für Operations Research
RWTH Aachen University



@mluebbecke

OR 2017 · Berlin · September 6, 2017