Mathematical Optimization + ML: Featuring Survey Insights From Forrester

GUROBI OPTIMIZATION

The World's Fastest Solver

Forrester®

Your Perfect AI Tech Team

Guest Speaker:
Mike Gualtieri, VP and Principal Analyst

September 17, 2019 – Gurobi Webinar

♡AI

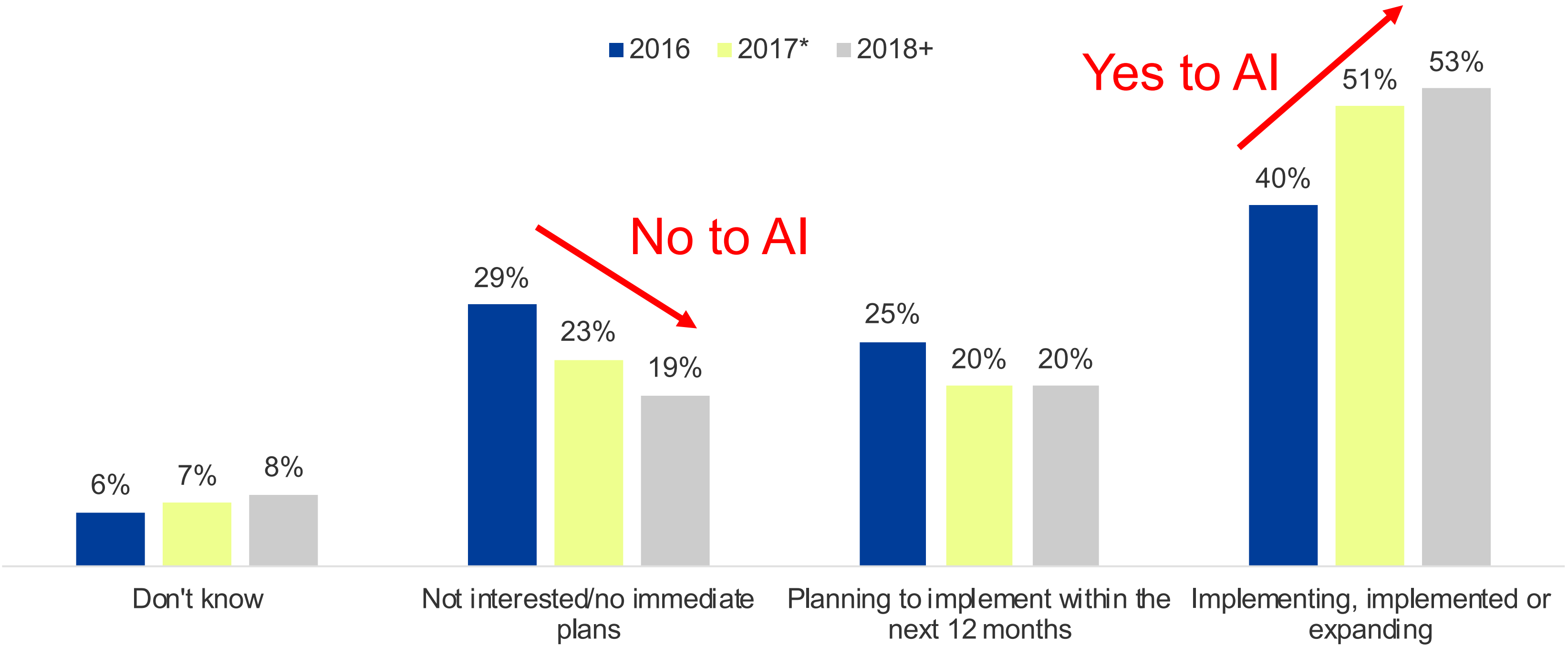

AI is a force for good.


It will make the world safer, healthcare more accessible, education personalized, manufacturing efficient, and will touch virtually every other aspect of humanity in net positive ways.


Enterprises must prioritize AI in order to be leaders in their industry.

Forrester projects that nearly every enterprise will use AI in five years.

“What are your firm's plans to use the following analytics technologies? (artificial intelligence)”

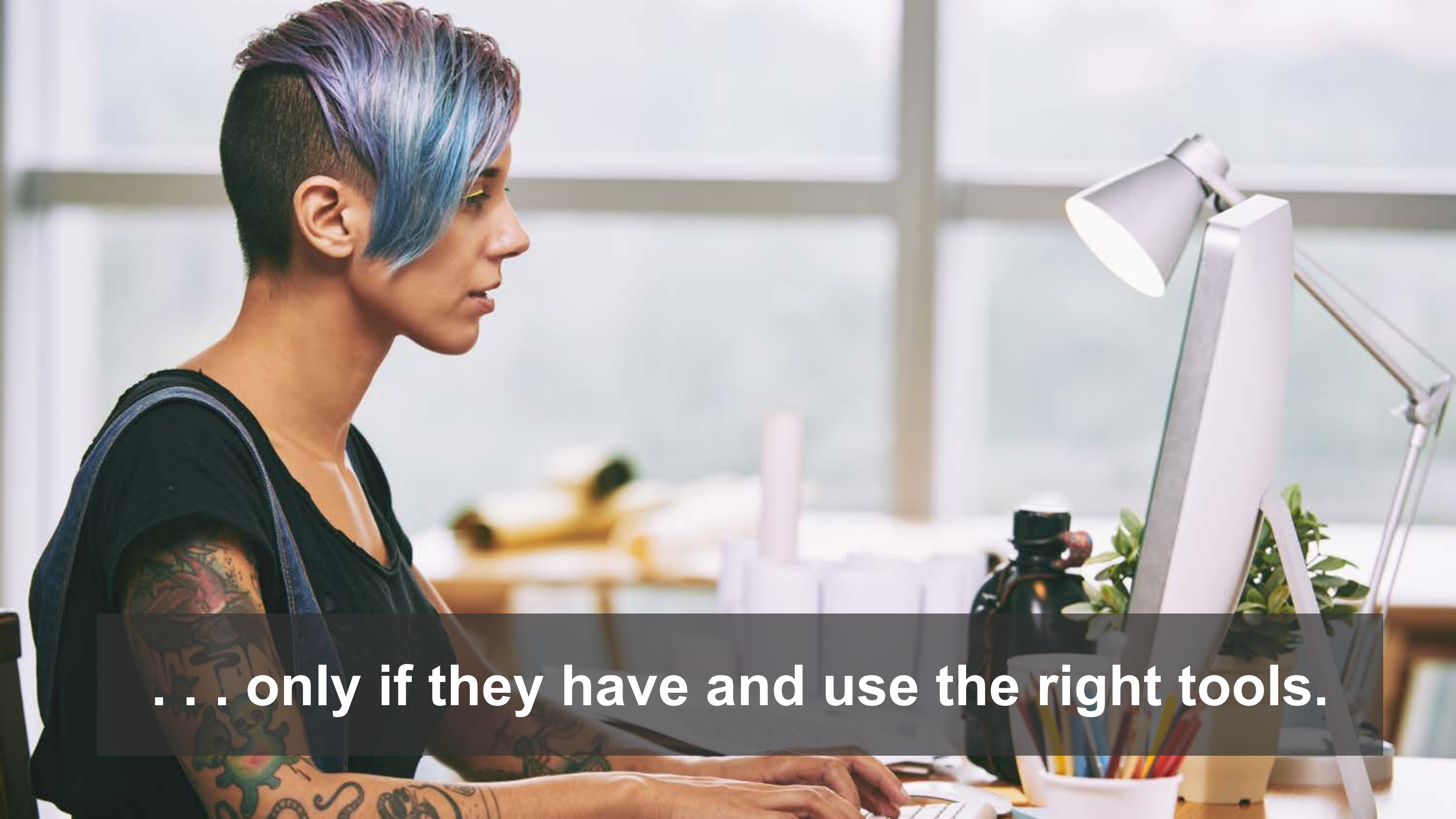


Base: 2,094, 2,106*, 1,742 + data and analytics decision makers
Source: Forrester Analytics Global Business Technographics® Data And Analytics Survey, 2016, 2017, 2018

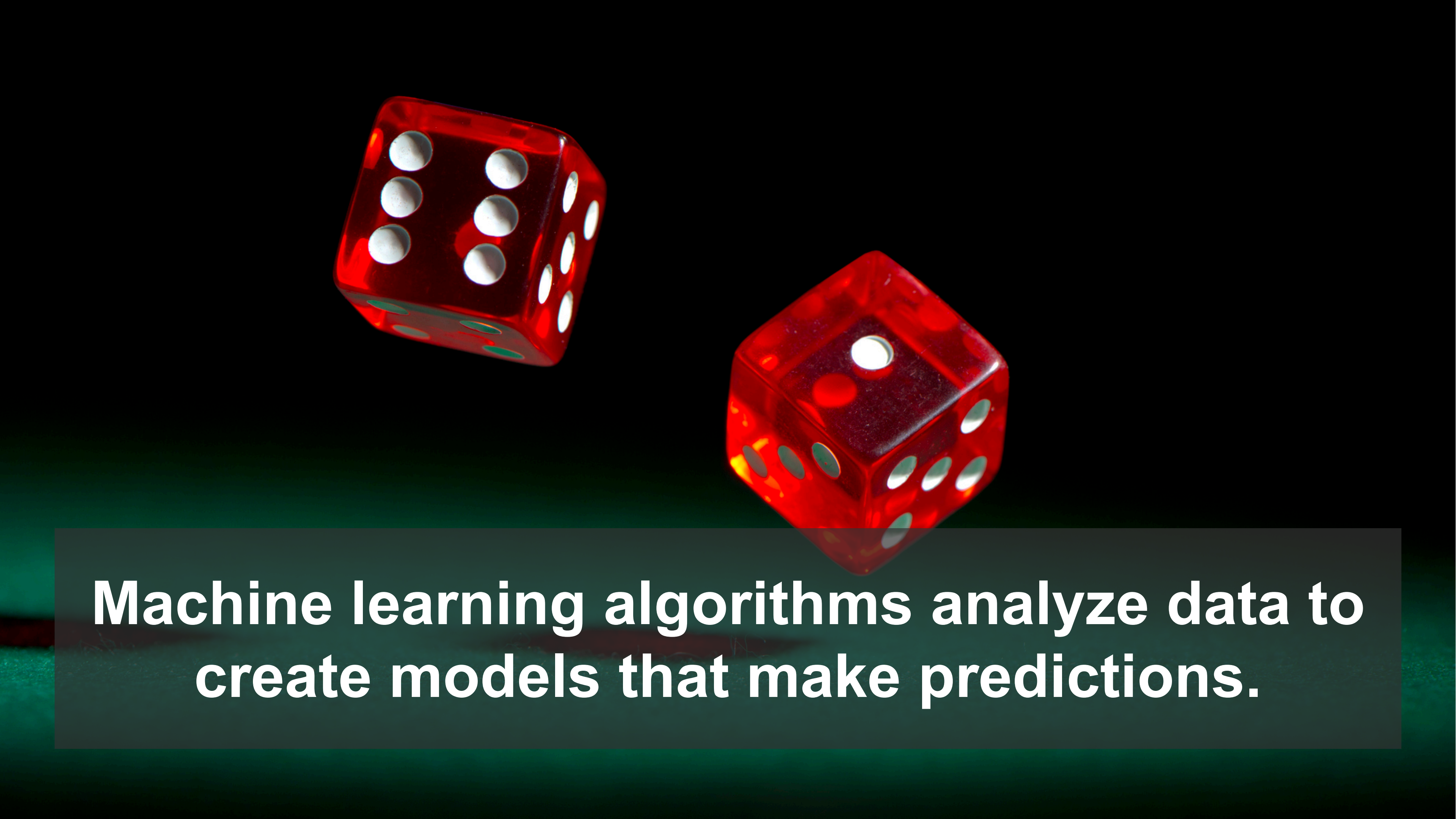

Artificial intelligence is real and ready.


There are as many use cases as there are
business processes and customer experiences.


Data scientists can make it happen . . .


. . . only if they have and use the right tools.

ML


Machine learning algorithms analyze data to create models that make predictions.

Machine learning (ML) algorithms *train* a *model* that takes inputs to make a prediction.

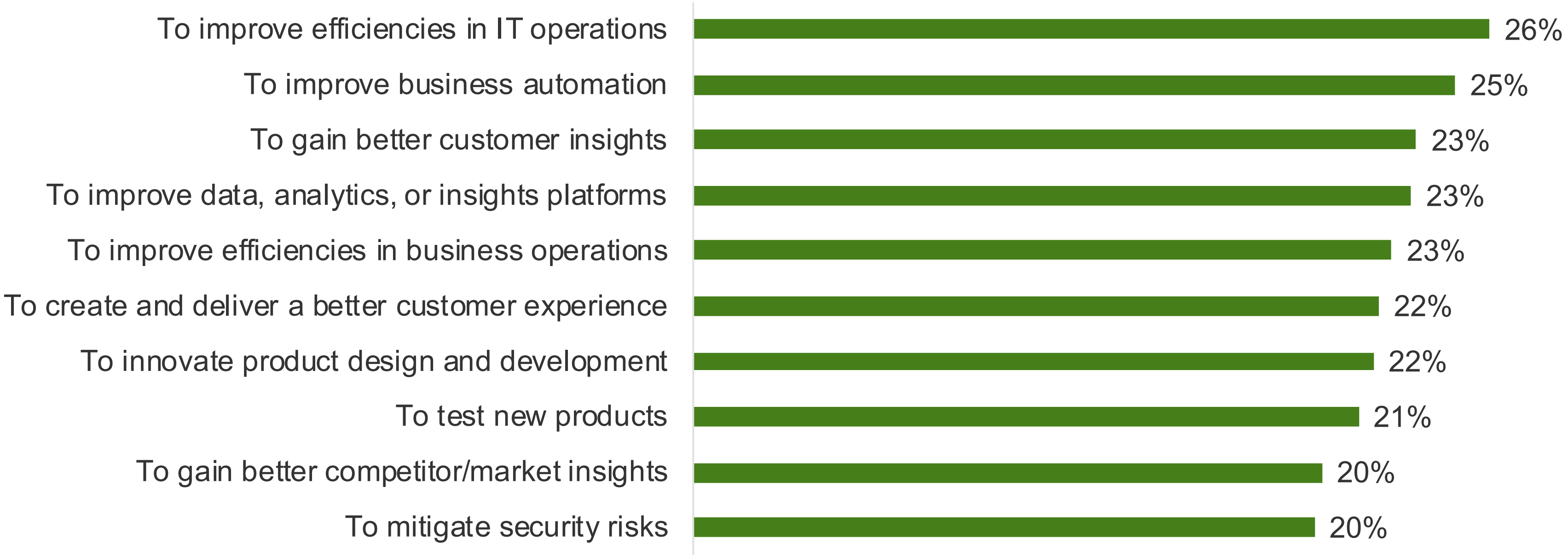
$$p = \text{model}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{x'}, \mathbf{y'}, \ldots)$$

Prediction
called *scoring* or *inferencing*

Machine learning model
generated by ML algorithm by analyzing data

Input variables
selected by the ML algorithm

"Which of the following use cases/application scenarios is your firm planning to use/currently using AI technologies for? (Top 10 responses shown)"



Base: 2,886 data and analytics decision makers whose firm is adopting AI
Note: Not all responses shown.
Source: Forrester Analytics Business Technographics Global Data And Analytics Survey, 2019

#UseCases


Predict supply-chain issues while there is still time to remediate now.

Predict who will launch what cyberattack before it happens.


Predict experiments that are more likely to prove the hypothesis to avoid wasting time.


Predict imminent machine failure.


Predict benefits eligibility fraud.


Predict the needs of infrastructure maintenance right now.


Predict price movements to find investment opportunities before the market does.

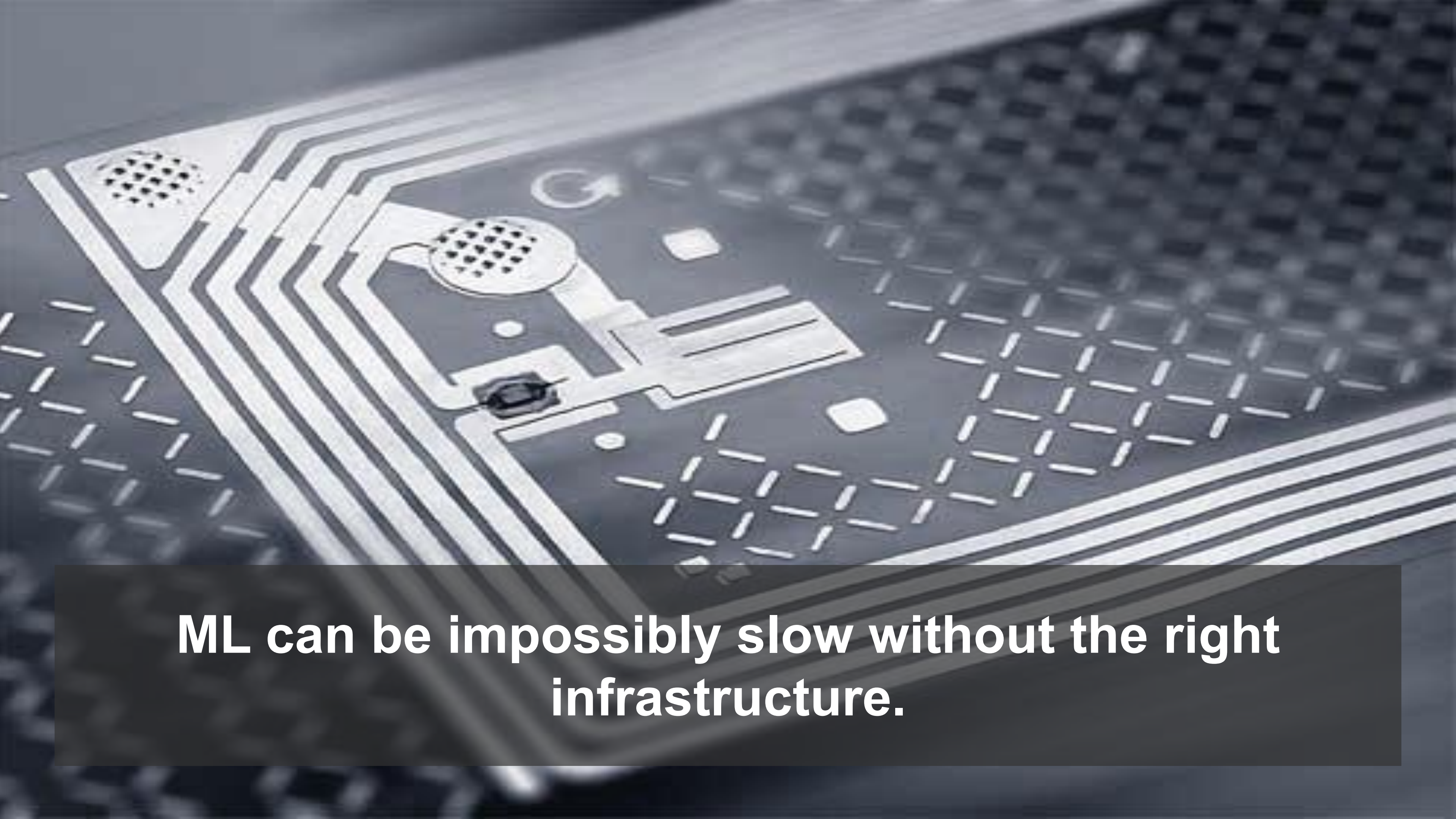
Predict customer propensity to buy more with targeted offers.

Image source: iStockphoto


Machine learning (ML) is not without challenges.


Business problems must be translated into a set of predictions.

Garbage in = garbage out.



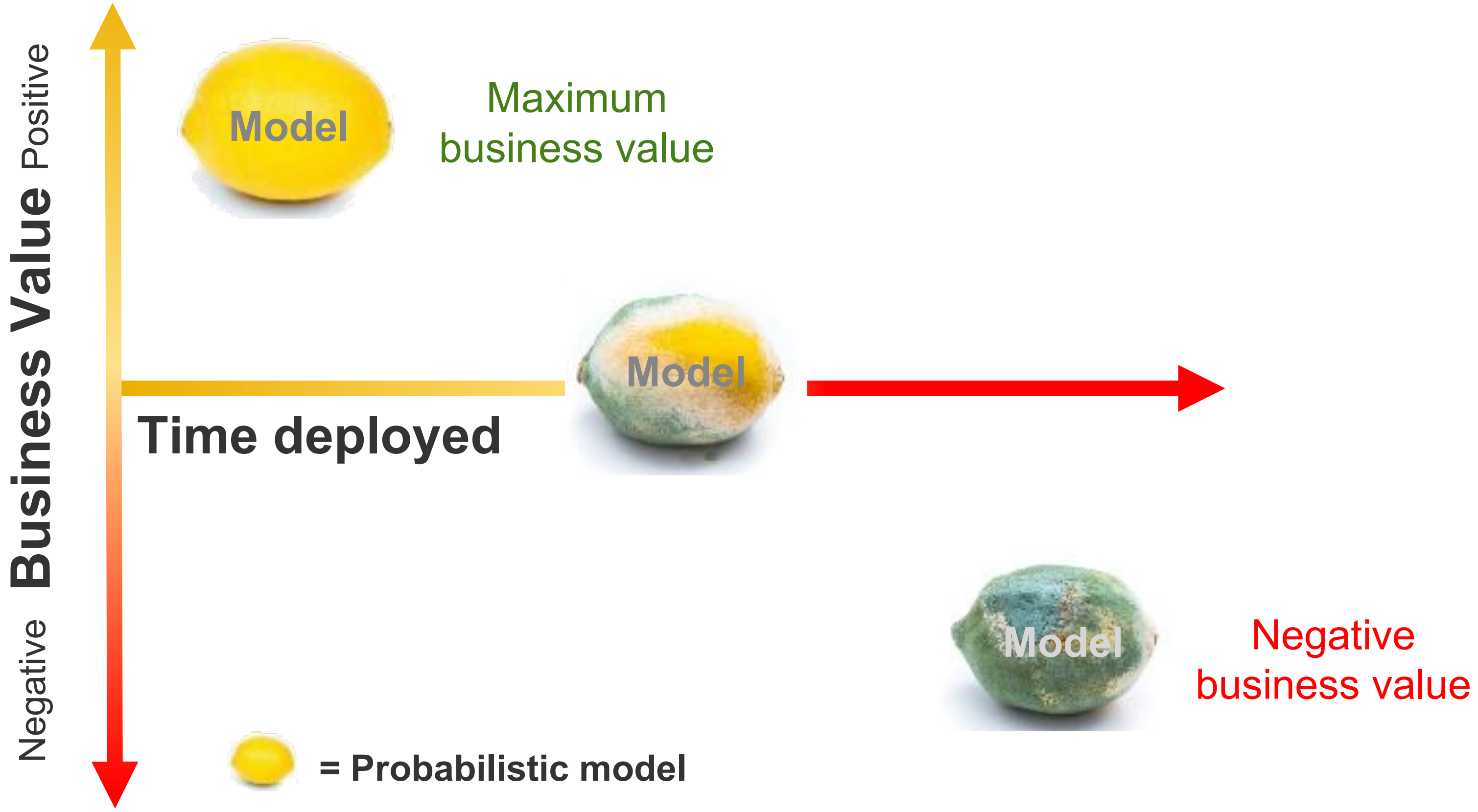
ML can be impossibly slow without the right infrastructure.

ML models can be very powerful and profitable, but understand that:

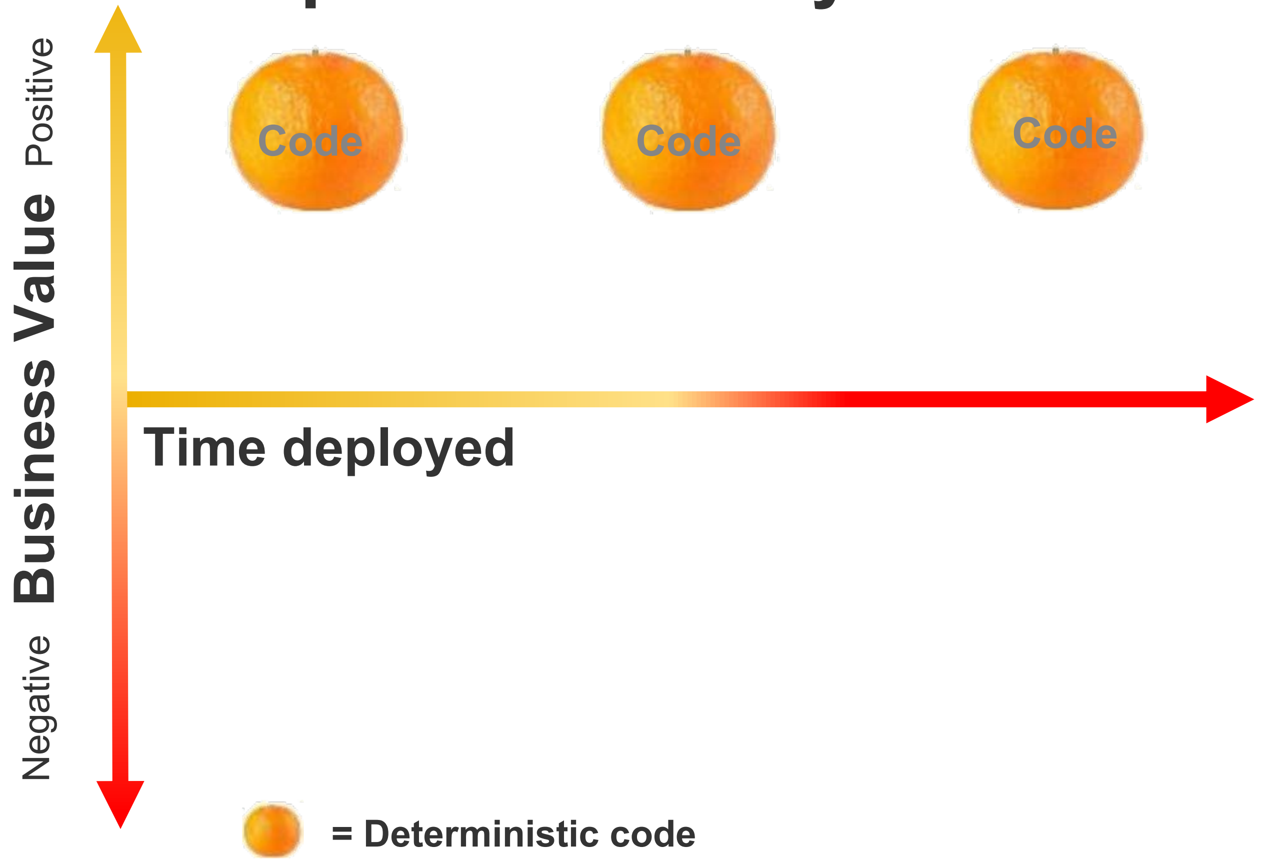
- Models are about probabilities, NOT absolutes.
  - E.g., 78% chance you will enjoy watching *Money Heist* on Netflix.
- Accurate models may NOT exist for every question.
  - E.g., elections, economic indicators, fashion, etc.
- ML models are based on correlation and probably, they are NOT causative.


ML models are like us; they must learn from experience.

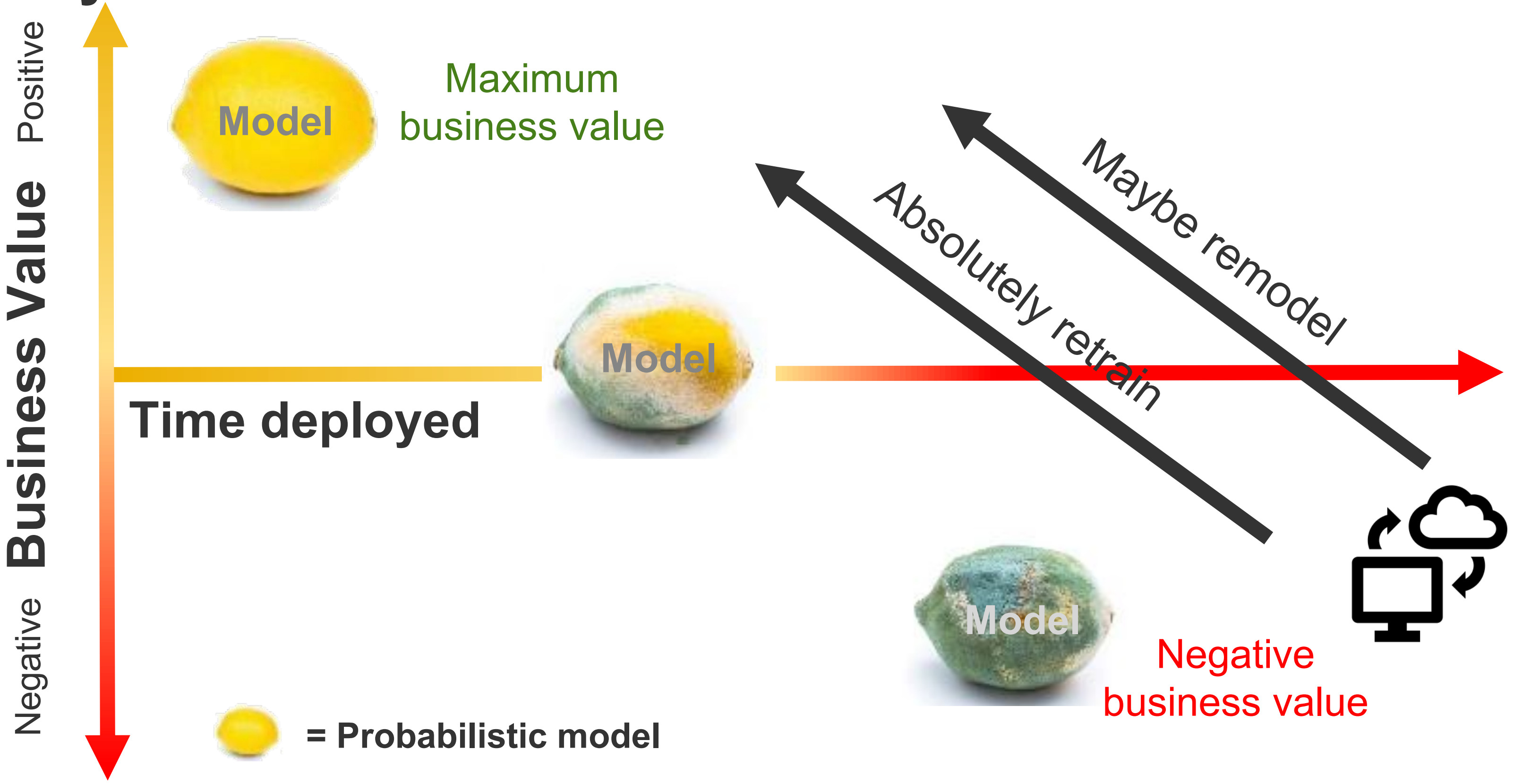
ML model performance can decay over time.



Dev-developed code always runs as written . . .



. . . but, ML models must be retrained on newer data to stay fresh.



♡MO


Mathematical optimization (MO) determines the best decision based on real-world constraints.

Mathematical optimization uses a *solver* to *calculate* the decision based on constraints.

$$[d] = \text{solver}(o(), \mathbf{c}^1(), \mathbf{c}^2(), \mathbf{c}^3(), \ldots)$$

Decisions
optimal input variables that constitute the best decision

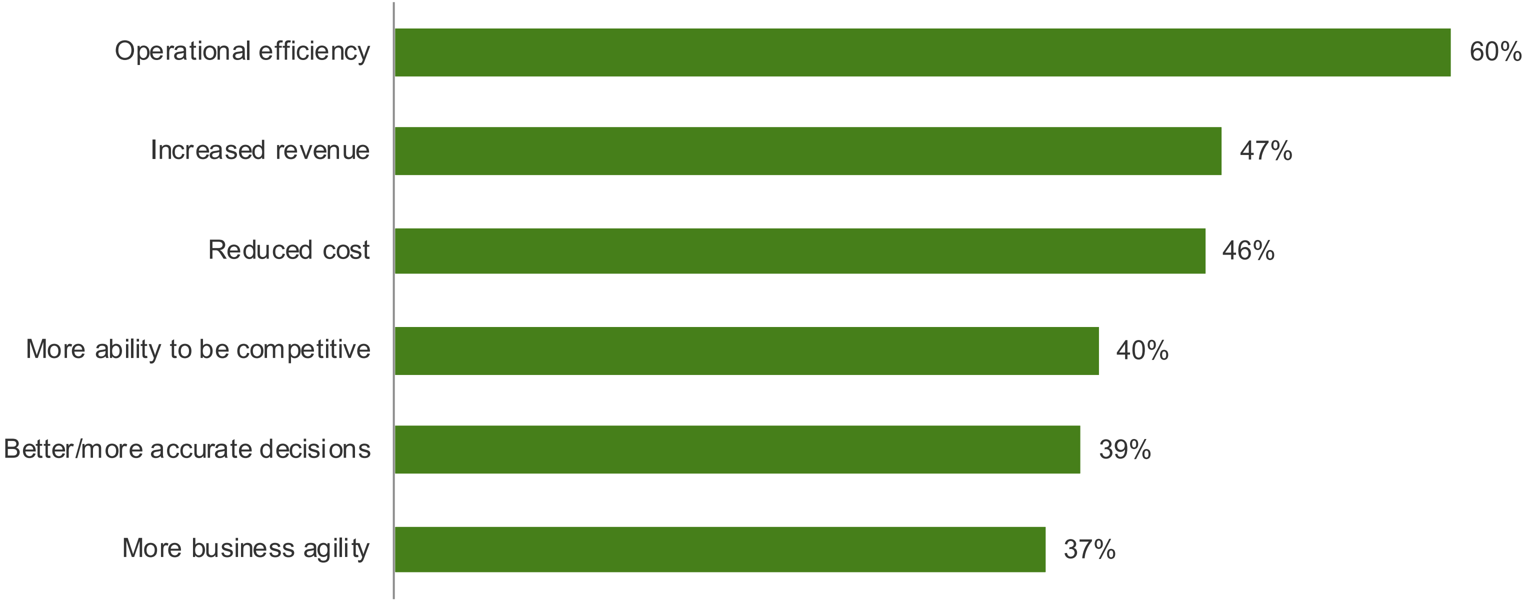
Mathematical optimizer
software the calculates the best possible decision

Objective function
defines a min or max that constitutes the best decision

Constraint functions
defined by business requirements

Mathematical optimization drives improvements across the enterprise.

"Which of the following benefits has your organization realized/do you expect to realize as a result of applying mathematical optimization for tasks like scheduling, sourcing, route planning, resource optimization, etc.? (Select all that apply)"



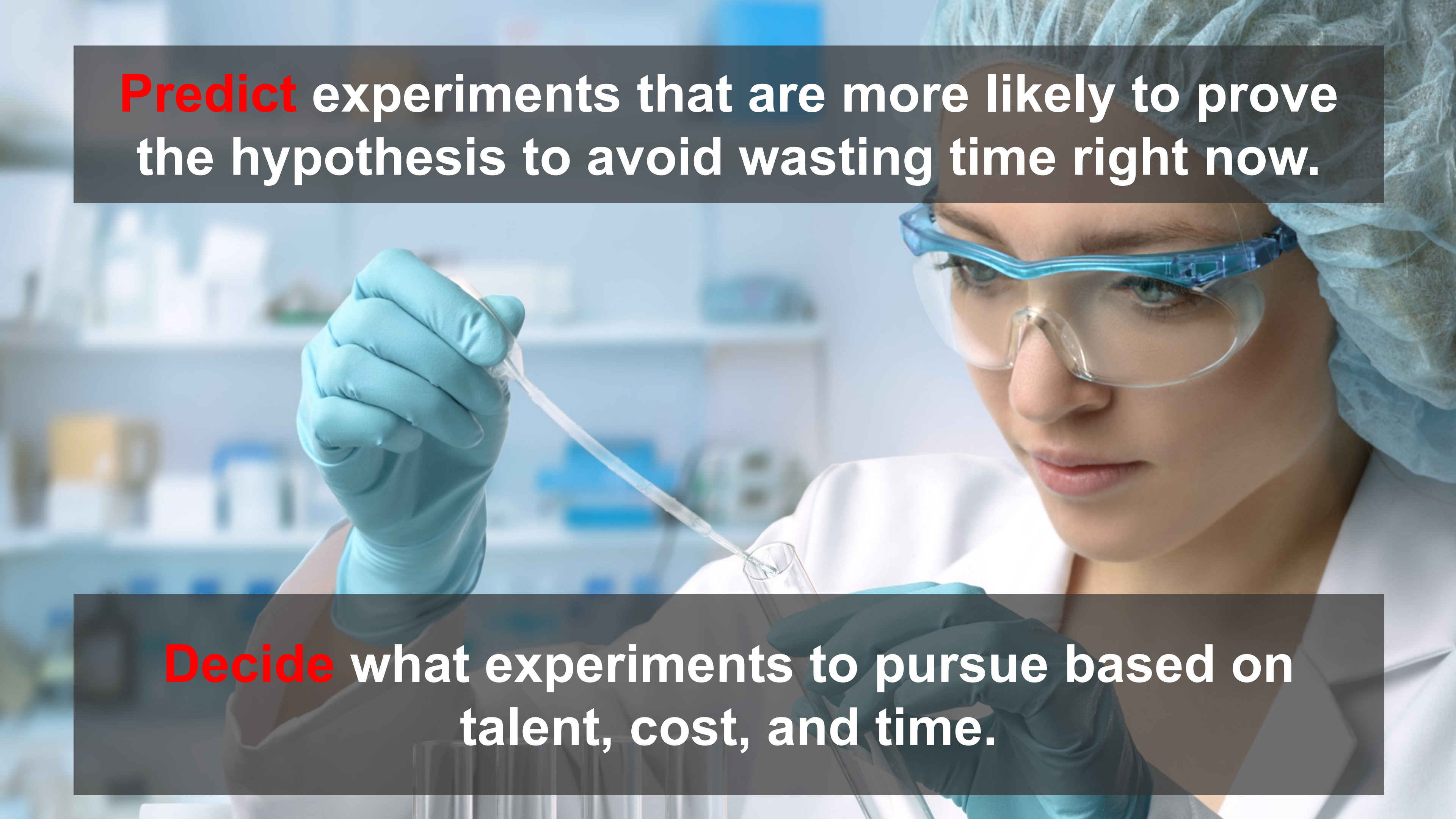
Base: 153 US +managers who are responsible for or influence their organizations' data science or execution strategy
Source: A commissioned study conducted by Forrester Consulting on behalf of Gurobi, July 2019

#UseCases


Predict supply-chain issues while there is still time to remediate now.
Decide the least costly way to reroute shipments.

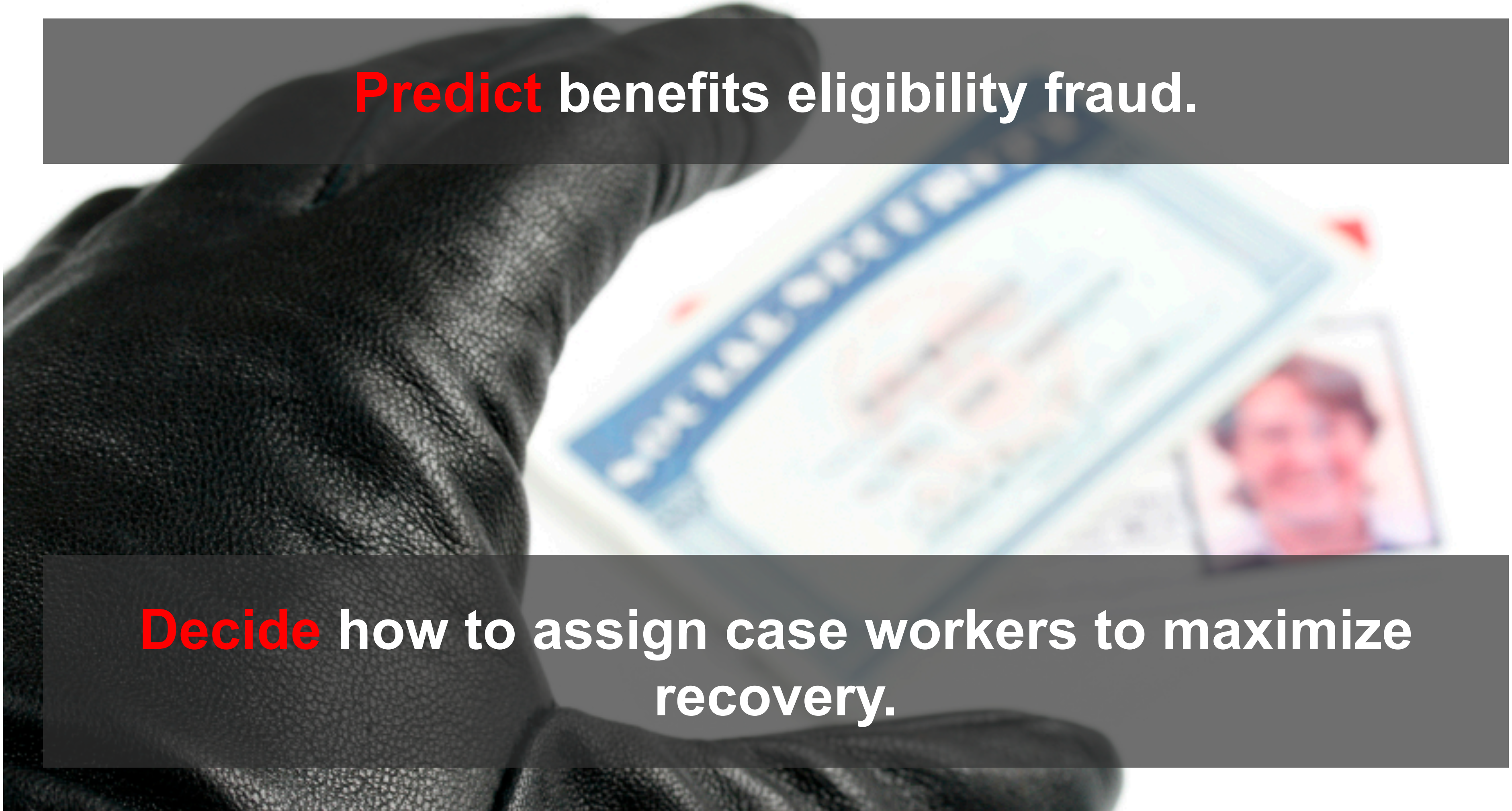
Predict who will launch what cyberattack before it happens.

Decide what investigators to assign to potential cyber threats based on investigator skill and potential damage.


Predict experiments that are more likely to prove the hypothesis to avoid wasting time right now.
Decide what experiments to pursue based on talent, cost, and time.

Predict imminent machine failure.

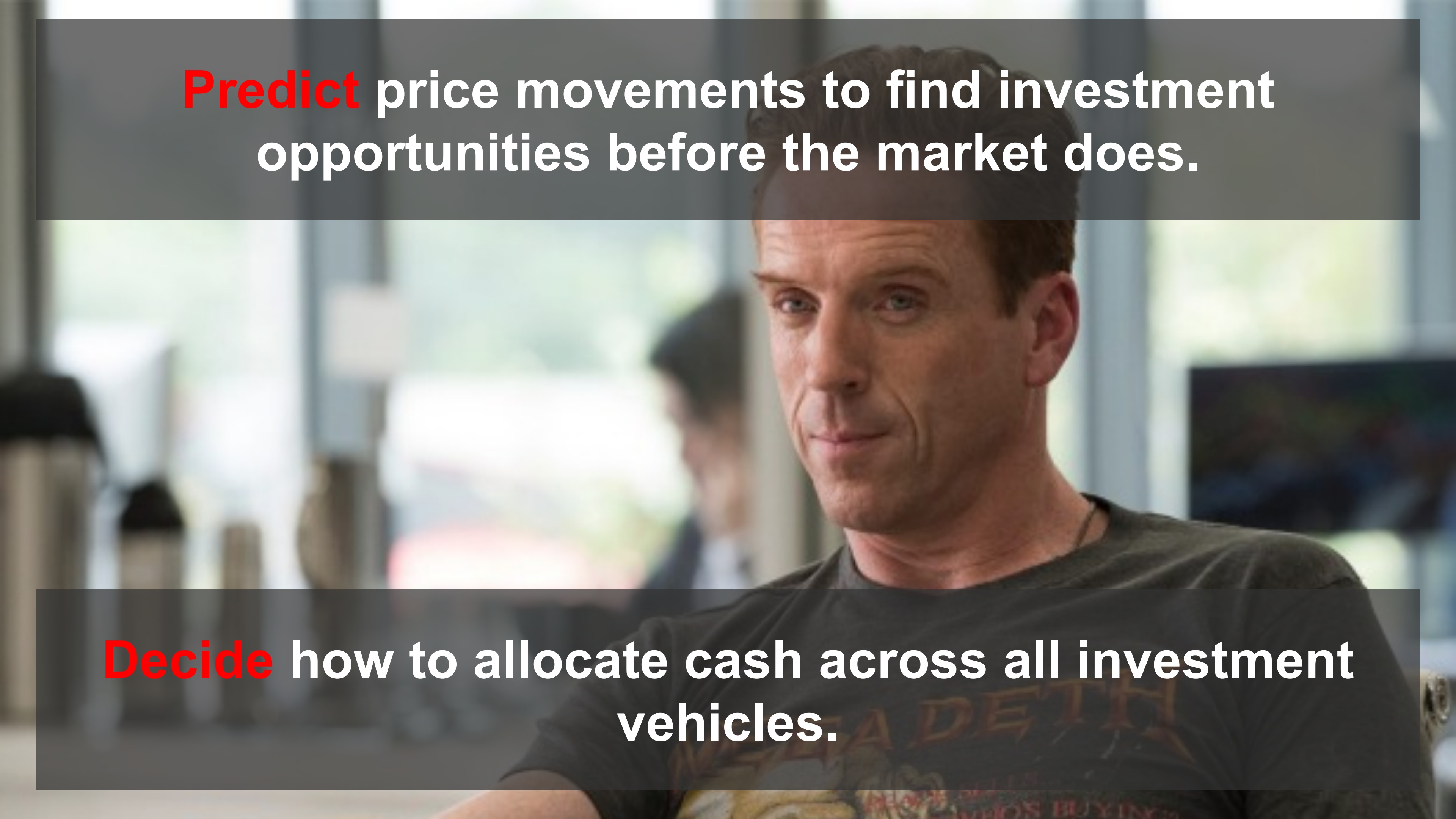
Decide when to shut the production line down to perform maintenance to minimize cost and customer complaints.



Predict benefits eligibility fraud.

Decide how to assign case workers to maximize recovery.

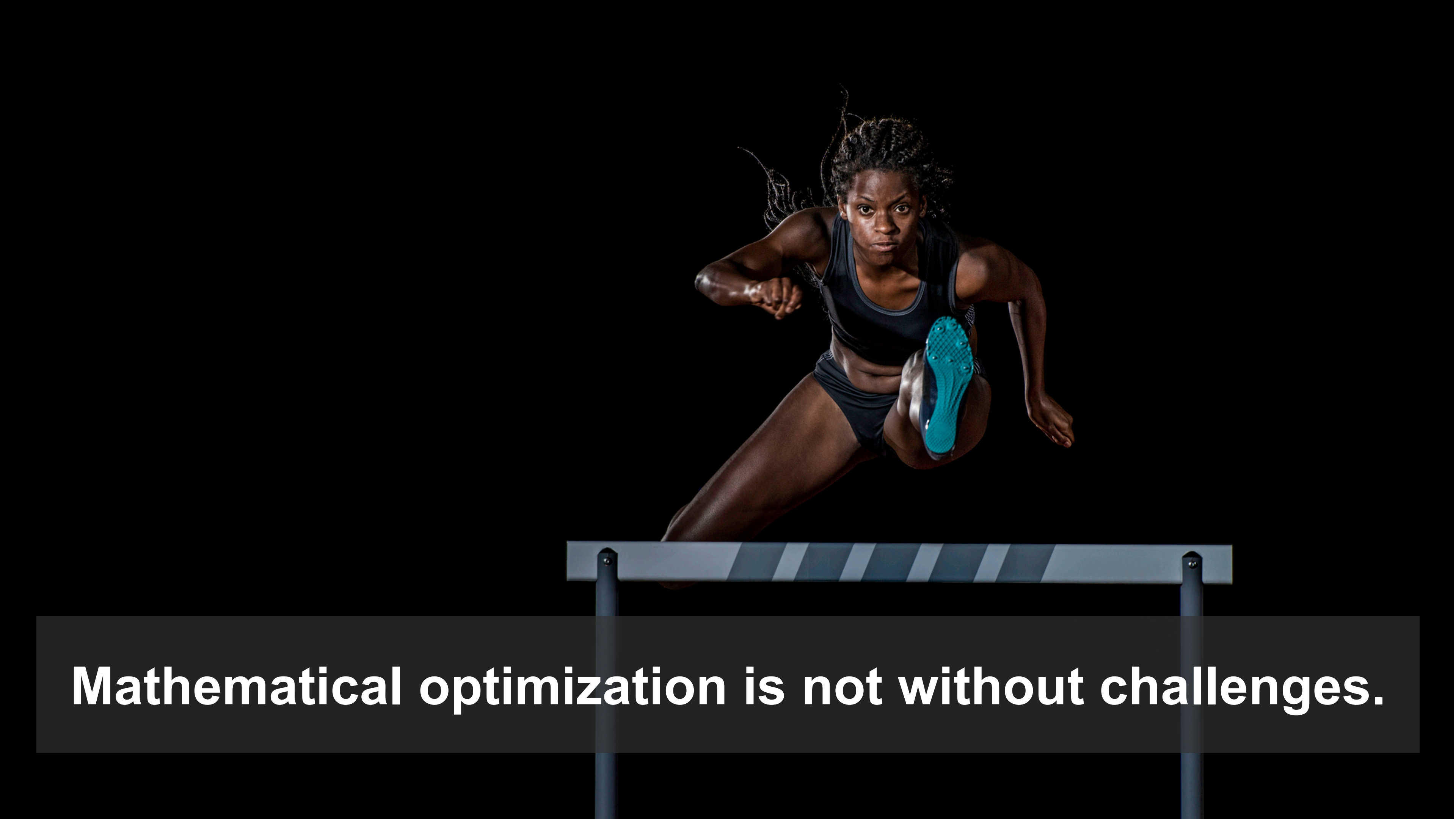

Predict the needs of infrastructure maintenance right now.
Decide how to assign maintenance teams based on cost and skill.


Predict price movements to find investment opportunities before the market does.
Decide how to allocate cash across all investment vehicles.

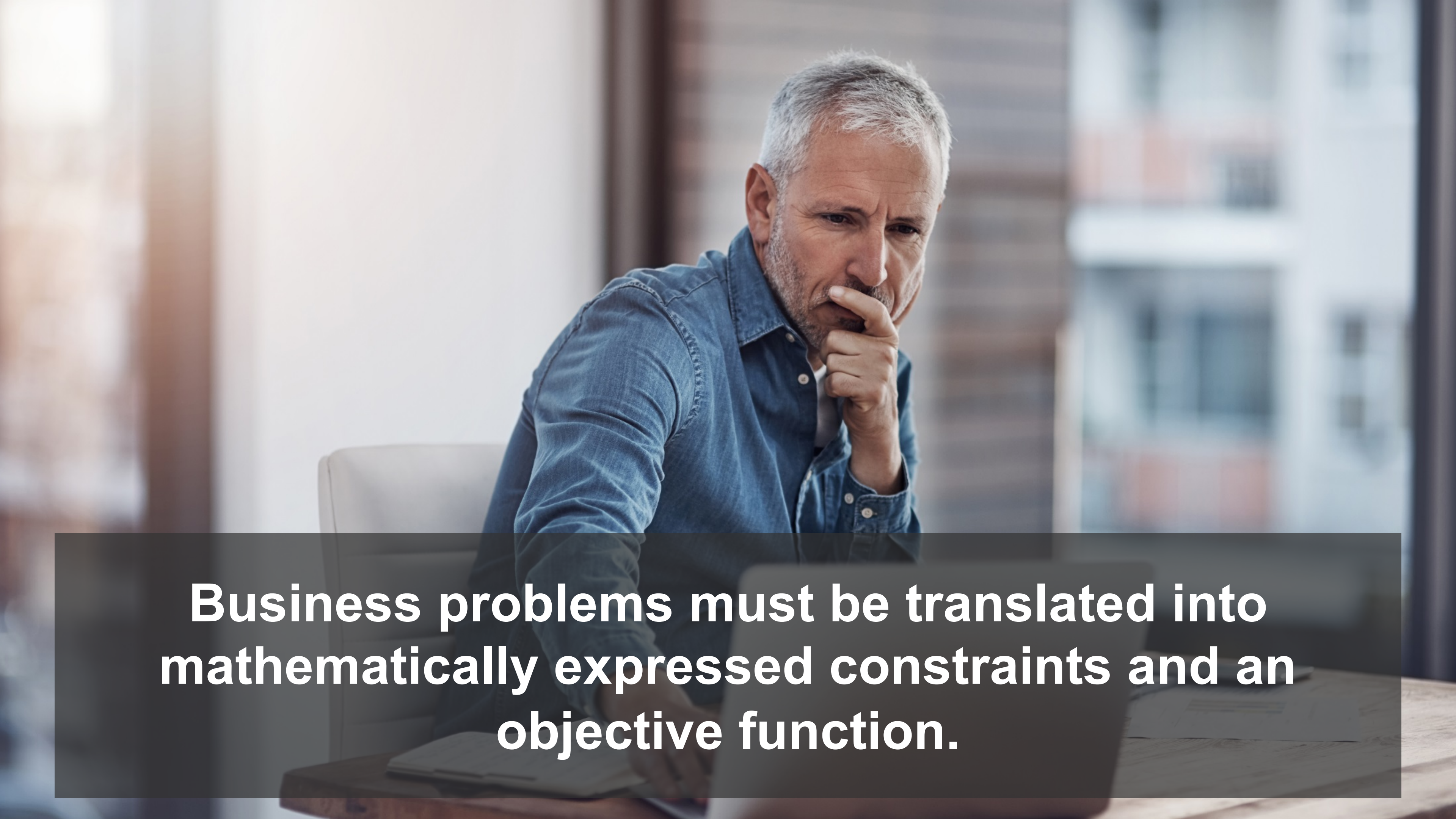
Predict customer propensity to buy more with targeted offers.

Decide how many discount coupons to offer to maximize revenue or to maximize profit.

Image source: iStockphoto

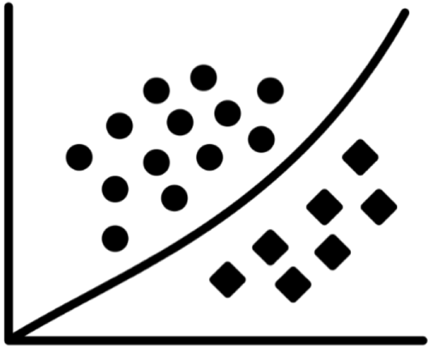


Mathematical optimization is not without challenges.

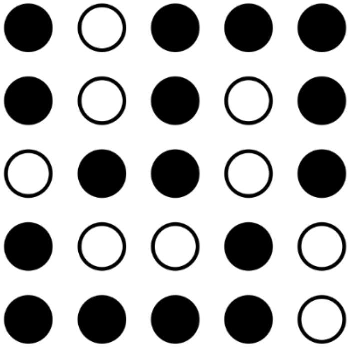

Business problems must be translated into mathematically expressed constraints and an objective function.


Garbage in = garbage out

Mathematical optimization (MO) can be impossibly slow, or impossible without performant solver software.



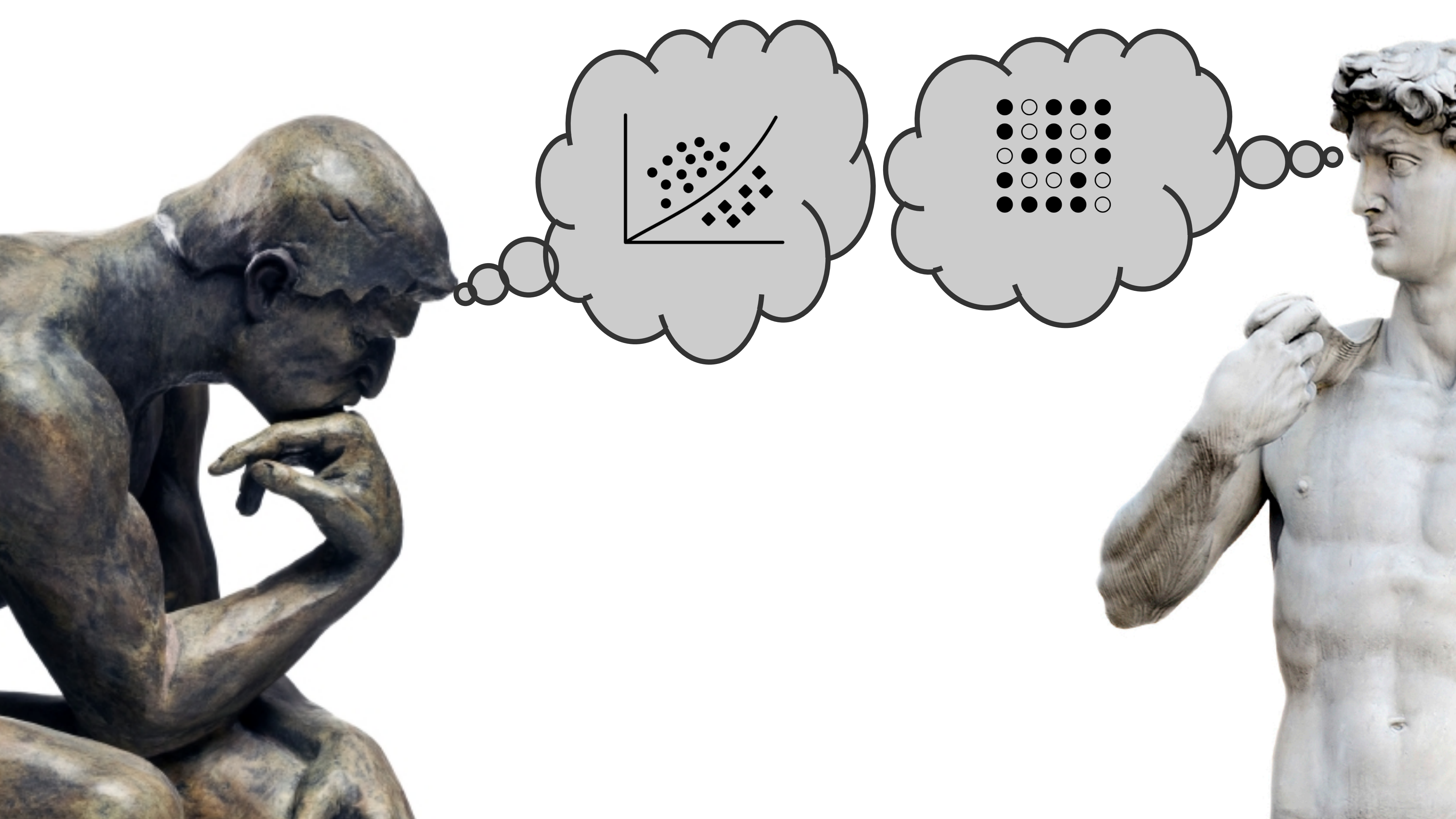
+




ML

+


MO

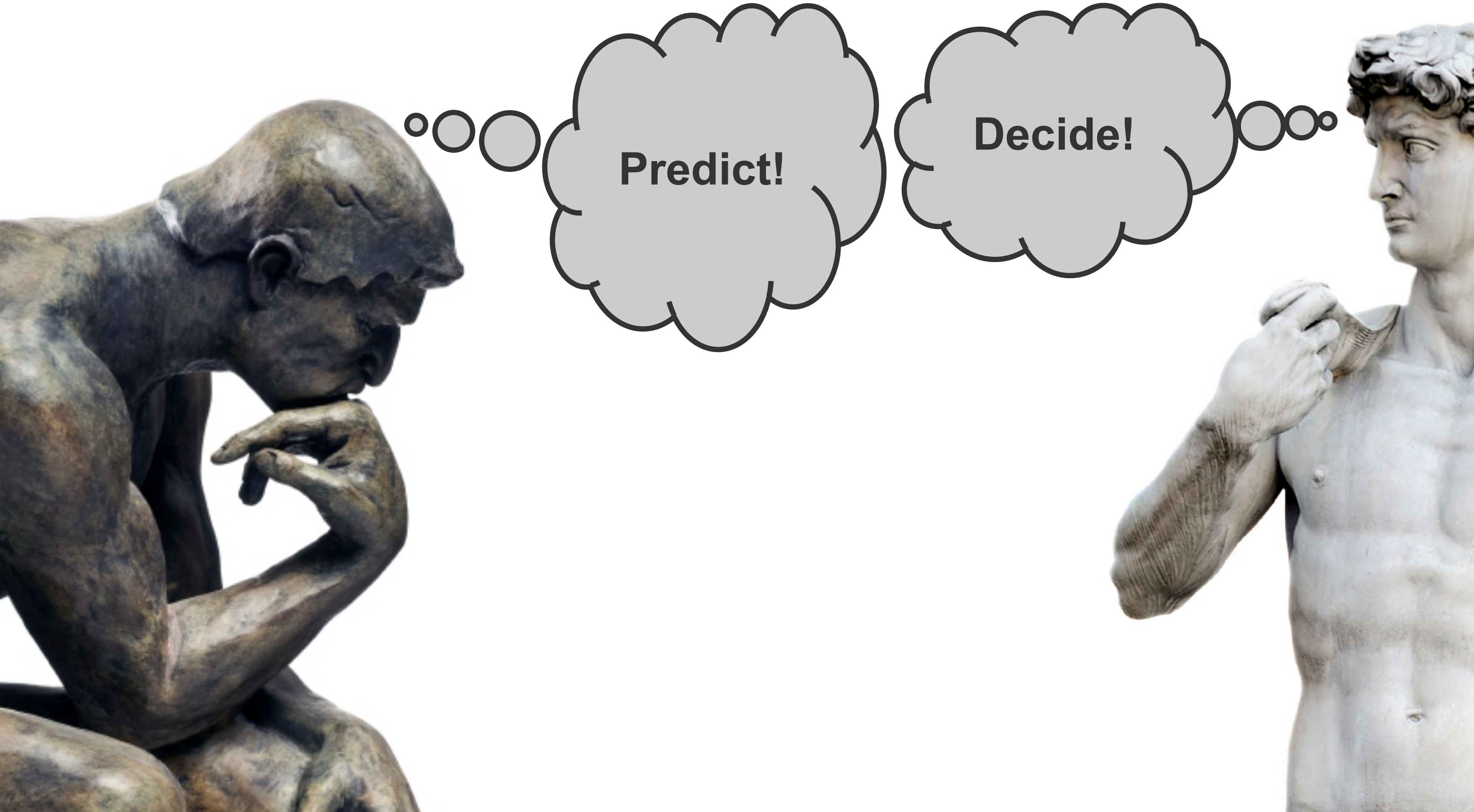


ML predictions can determine the need to make a MO decision.

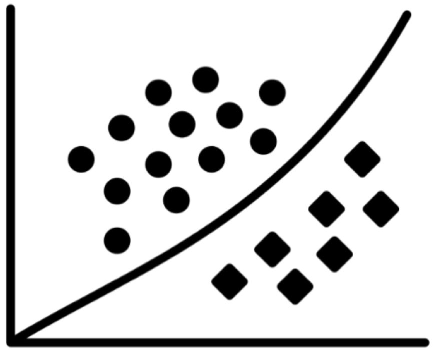
ML predictions can be used as MO decision constraints.

MO decisions can be used as ML model training features.

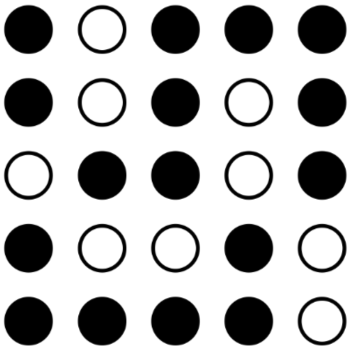
MO decisions can be used as ML model scoring/inferencing inputs.


Predict!
Decide!

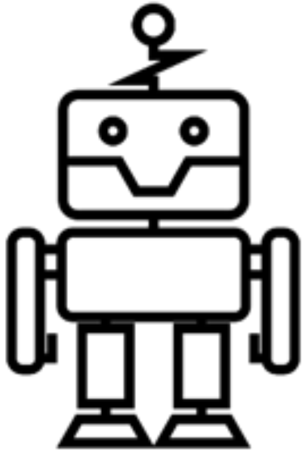
♡ AI



+



=



ML + MO = AI

FORRESTER®

Mike Gualtieri
mgualtieri@forrester.com

Thank you ♡

FORRESTER.COM

Mathematical Optimization: A Closer Look

Ed Rothberg, CEO and Co-Founder

GUROBI OPTIMIZATION

The World's Fastest Solver

Mathematical Optimization – A Closer Look

GUROBI OPTIMIZATION

- **Primary mathematical optimization technology**
  - Mixed Integer Programming (MIP)
- **Optimize an objective function over a set of decision variables subject to a set of constraints**
  - Constraints come from activities that compete for resources
    - Budget, machines, trucks, time slots, workers, etc.
  - Objective is the quantity you wish to maximize (or minimize)
    - Maximize profit, minimize waste, minimize late orders, etc.
- **Number of possible solutions is typically astronomical**
  - Billions, trillions barely begins to capture it
- **Optimization combs through possible solutions to find those that maximize your objective**
  - Always gives you a measure of the quality of the solution
  - If you let it run long enough, it gives you the optimal solution

Optimization and Machine Learning – Enabled by Data

- Similar histories for optimization and ML
- Enabled by explosion of available data and computing resources
- Fundamental techniques are 50+ years old in both cases
- Big improvements in underlying optimization technology over time
  - Measured in factors of millions

Optimization and ML – Broad Applicability

- **Both approaches have broad applicability across a number of industries**
  - Optimization used in over 40 different industries
  - New ones popping up all the time
- **Optimization less visible, mainly because applications are usually on a larger scale**
  - Typically used to make sets of complex, inter-related decisions
  - Not as visible as showcase ML applications like speech recognition or image recognition or autonomous vehicles

Customer Applications of MIP (2011-2012)

- Accounting
- Advertising
- Agriculture
- Airlines
- ATM provisioning
- Compilers
- Defense
- Electrical power
- Energy
- Finance
- Food service
- Forestry
- Gas distribution
- Government
- Internet applications
- Logistics/supply chain
- Medical
- Mining
- National research labs
- Online dating
- Portfolio management
- Railways
- Recycling
- Revenue management
- Semiconductor
- Shipping
- Social networking
- Sourcing
- Sports betting
- Sports scheduling
- Statistics
- Steel Manufacturing
- Telecommunications
- Transportation
- Utilities
- Workforce Management

Sampling of Gurobi customers

Industries Transformed by MIP – Airlines

- **Perfect industry for applying optimization**
  - Many complex, high-stakes decisions
- **One of the earliest large-scale adopters of optimization**
  - Adoption started in the 1970's
  - Nearly every aspect of operating an airline is influenced by an optimization model

Industries Transformed by MIP – Supply Chain

- **In the 1980's, software dominated by rules of thumb**
  - Example: theory of constraints (*The Goal*, Goldratt)
- **MIP widely adopted in the 1990's**
- **Now the standard technology for supply-chain planning**
  - SAP, Oracle, JDA, Manhattan Associates, …

Industries Transformed by MIP – Electrical Power

- **Electrical power deregulated in the late 1990's**
- **Need to create a market for electricity**
- **Early solution techniques:**
  - Heuristics (Lagrangean relaxation)
  - MIP (lots of models; no real usage)
- **EPRI report, June 1989:**
  - "Mixed-integer programming (MIP) is a powerful modeling tool. 'They are, however, theoretically complicated and computationally cumbersome'"
- **DIMACS meeting 1999:**
  - Bob Bixby demonstrated that MIP had improved to the point where practical power models could be solved
- **Within a few years, nearly every grid operator in the world was using MIP to solve these models**

Industries Transformed by MIP – Sports Scheduling

- **Computing sports schedules quite complicated**
  - Stadium constraints, travel constraints, TV schedules, …
- **Done by hand for decades**
  - Example: Henry and Holly Stephenson scheduled Major League Baseball "by hand" from 1981-2004
- **Schedules now done using MIP:**
  - MLB since 2004
  - NFL since 2007

Combining Machine Learning and Optimization

- **Not direct competitors – solve different problems**

- **Several ways in which they can work together**
  - Machine learning model feeds optimization model
    - ML makes predictions, optimization recommends actions
  - Optimization model feeds machine learning model
    - Optimization finds the best actions to take for different inputs
    - Machine learning model learns relationships between inputs and best actions
  - Tight integration of technologies

Machine Learning Feeding An Optimization Model

- **Example: Markdown Optimization and Blue Yonder**
- **Objective:**
  - Determine best schedule for marking down prices to clear out old inventory
- **Inputs:**
  - Retailers stock multiple items in multiple stores
  - Customer reaction to markdown depends on item and store
- **Constraints:**
  - Relationships between prices of related items
    - Don't sell 2L for less than 1L
  - Workload limits
    - Not too many price changes per day

Machine Learning Feeding An Optimization Model

- **Blue Yonder GmbH**
  - Cloud-based Predictive Applications for retail
- **Machine learning model**
  - Use historical data to predict product sales volume under different conditions
    - Different prices
    - Different stores
- **Optimization model**
  - Use sales volume predictions to compute optimal markdown schedules that satisfy constraints
- **Demonstrated 5% increase in revenue during markdown phase**

Tight Integration – Simple Example

- A new look at a fundamental problem in statistics
  - Best subset regression
- Trivial to state as a MIP
  - “MIP is NP-hard – we have to use heuristics”
- 30 years of heuristics
- Effective, but…
  - No quality guarantees
  - Falls apart in presence of side constraints
- Modern MIP solver can find optimal solutions to large problems in seconds

GUROBI OPTIMIZATION

The Annals of Statistics

Info | Current issue | All issues | Search

Ann. Statist.
Volume 44, Number 2 (2016), 813-852.

← Previous article | TOC | Next article →

Best subset selection via a modern optimization lens

Dimitris Bertsimas, Angela King, and Rahul Mazumder

Full-text: Open access

Enhanced PDF (788 KB)

Abstract | Article info and citation | First page | References | Supplemental materials

Abstract

In the period 1991–2015, algorithmic advances in Mixed Integer Optimization (MIO) coupled with hardware improvements have resulted in an astonishing 450 billion factor speedup in solving MIO problems. We present a MIO approach for solving the classical best subset selection problem of choosing k out of p features in linear regression given n observations. We develop a discrete extension of modern first-order continuous optimization methods to find high quality feasible solutions that we use as warm starts to a MIO solver that finds provably optimal solutions. The resulting algorithm (a) provides a solution with a guarantee on its suboptimality even if we terminate the algorithm early, (b) can accommodate side constraints on the coefficients of the linear regression and (c) extends to finding best subset solutions for the least absolute deviation loss function. Using a wide variety of synthetic and real datasets, we demonstrate that our approach solves problems with n in the 1000s and p in the 100s in minutes to provable optimality, and finds *near* optimal solutions for n in the 100s and p in the 1000s in minutes. We also establish via numerical experiments that the MIO approach performs better than Lasso and other popularly used sparse learning procedures, in terms of achieving sparse solutions with good predictive power.

Ingredients for Optimization Success

- A problem that involves multiple activities competing for scarce resources
- Available data that captures the current state and upcoming demands of these activities
- A well-formed objective function
- A data scientist that has the ability to systematically state the objective and constraints in a mathematical form

- Process of formulating an optimization model often as valuable as the model itself

Data Scientist Challenges

- Main challenge is to recognize which problems are ML and which are optimization
  - Often the answer is both
- Optimization opens up a much broader set of problems to data science
- Optimization should be a part of any data scientist's toolbox

Resources

- **Visit www.gurobi.com and select 'I am a Data Scientist.' or browse through our "Resources"**
  - You can find case studies, whitepapers, modeling examples, instructional videos and Optimization Application Demos to introduce you to MIP and help you get started.
- **Get a free 30-day trial of Gurobi**
  - www.gurobi.com/eval
- **Need help leveraging Optimization for your business?**
  - Contact us at info@gurobi.com

Thank You – Questions?

GUROBI OPTIMIZATION

The World's Fastest Solver