OliverWyman

UNLOCKING PREDICTIVE POWER WITH UNCONSTRAINED LOSS MODELS

Nickolas A. Alvarado, Brett Nunes

A business of Marsh McLennan

ACTUARIAL

CONFIDENTIALITY

Our clients' industries are extremely competitive, and the maintenance of confidentiality with respect to our clients' plans and data is critical. Oliver Wyman rigorously applies internal confidentiality practices to protect the confidentiality of all client information.

Similarly, our industry is very competitive. We view our approaches and insights as proprietary and therefore look to our clients to protect our interests in our proposals, presentations, methodologies, and analytical techniques. Under no circumstances should this material be shared with any third party without the prior written consent of Oliver Wyman.

DATA, DATA, DATA!

~70%

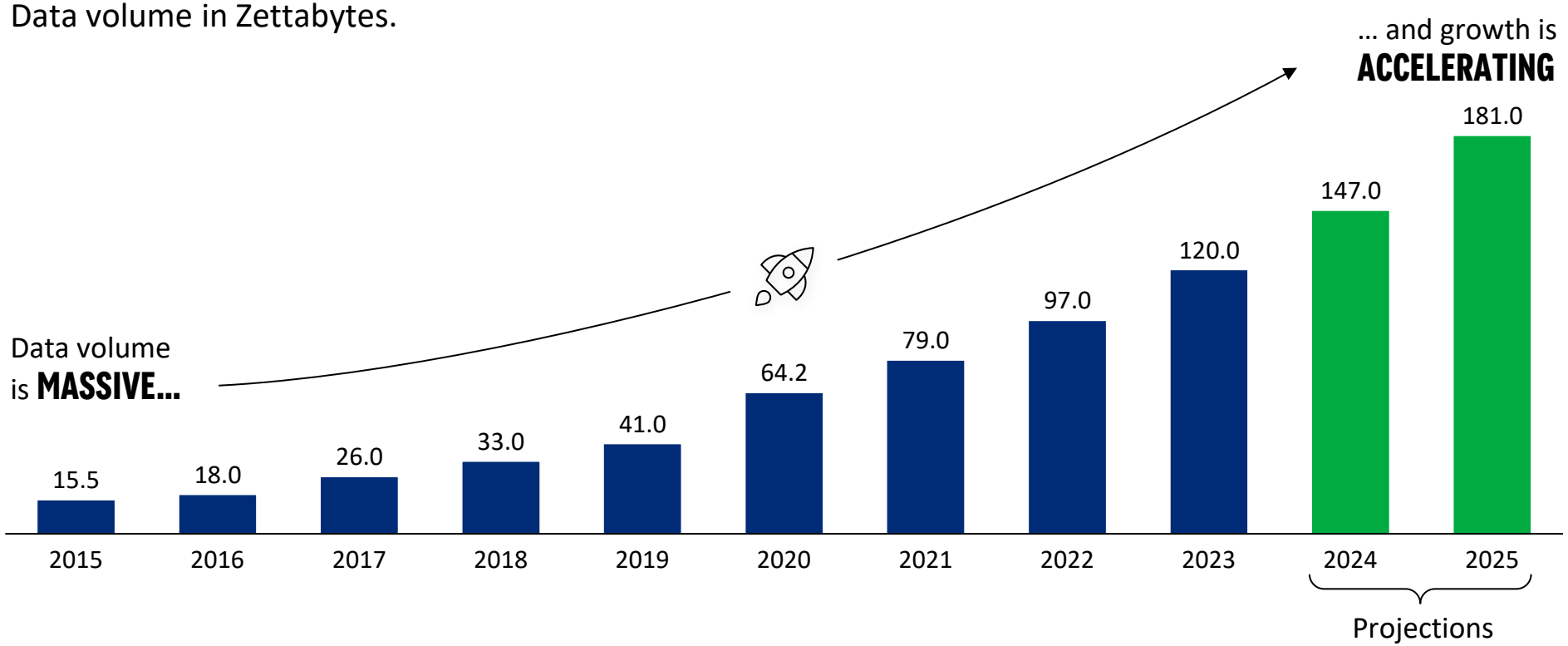
of data is user generated

~90%

of data is unstructured and grows faster than structured data

Data created captured, copied and consumed

Data volume in Zettabytes.



Data size scale

Prefix	Scale 10^x	Relatable Object
Kilo	10^3	Text documents (no images)
Mega	10^6	Digital photo
Giga	10^9	15 min to 1 hr video
Tera	10^12	500 to 1,000 hours of HD video

Prefix	Scale 10^x	Relatable Object
Peta	10^15	Netflix catalog size
Exa	10^18	Monthly internet traffic in 2020
Zetta	10^21	Netflix catalog size x 1 million
Yotta	10^24	Netflix catalog size x 1 billion

LIMITATIONS OF TRADITIONAL ALGORITHMS

Traditional models were designed in an era of limited data

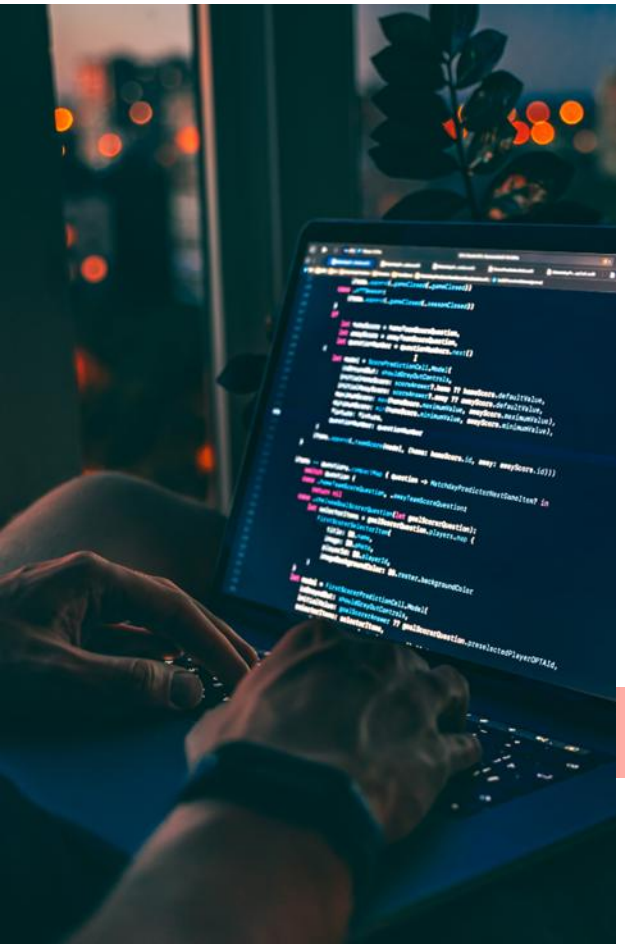
Limited data	Observation counts	Column counts	
Traditional models	Linear regression	GLMs	Actuarial reserving methods
Constraints	Business rules	Need for interpretability	Regulations

Constraints not necessarily bad

Model constraints allowed for creative solutions

- Automated variable selection (elastic net)
- Grid search (parameter optimization)
- Ensembles
- Random forests and GBMs
- Principal component analysis

CONSTRAINTS



Zero constraints are an **impossibility**.

Examples

- Data
- Resources
- Regulations/law
- Privacy
- Business
- Values and culture

CONSTRAINTS LIMIT BUSINESS INSIGHTS

UNCONSTRAINED MODELS

UNCONSTRAINED: **no limits on the data or the model/algorithm.**

MAXIMUM PREDICTIVE ACCURACY AND POWER **most impactful insights.**

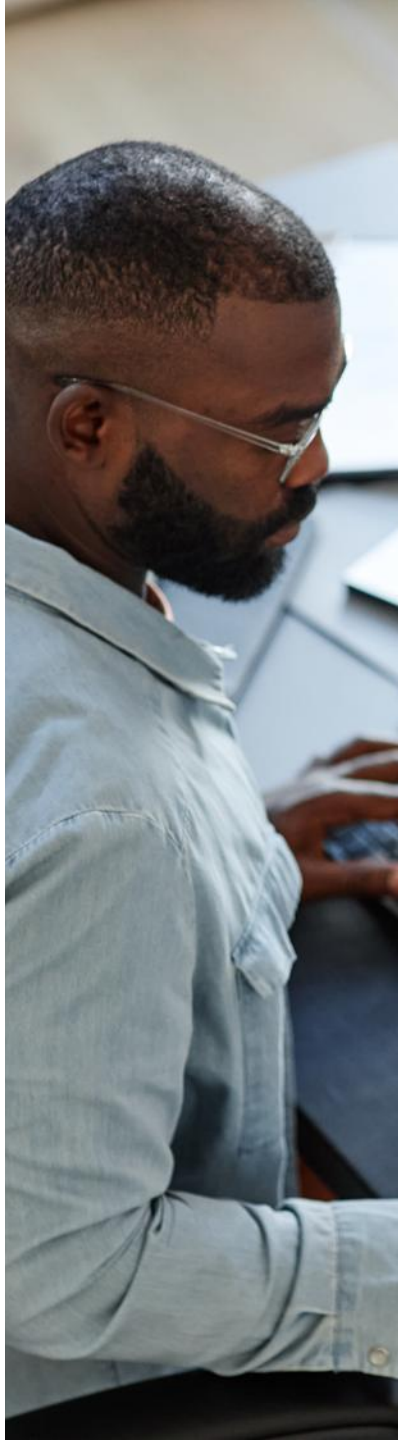
Strategic tool

Not replacements for traditional algorithms – they are used together.

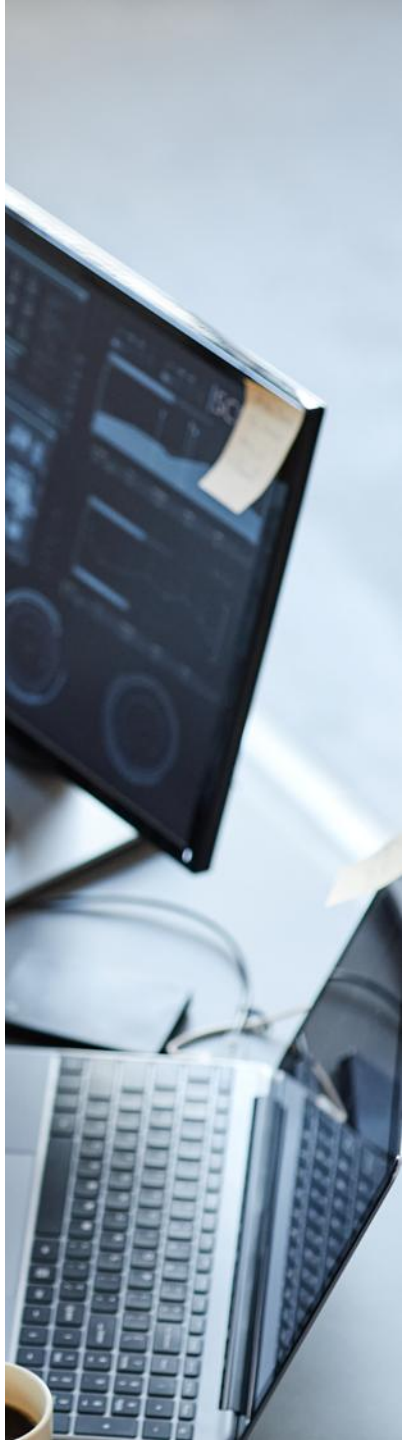
Leverage diverse data sources
- Structured (tabular) data (i.e., traditional data)
- Unstructured data (e.g., call transcripts, notes)
- External data (e.g., weather, vehicle traffic, satellite imagery, web traffic)

Advanced algorithms are leveraged (Gradient Boosted Machines, Neural Networks)
- Capable of finding complex relationships in data that traditional models miss or cannot handle
- Handle very large datasets

Requires skill and expertise
- Cannot throw everything into a blender and expect usable results
- Performance testing is critical



MINIMIZE GAP TO MAXIMIZE VALUE



GAP ANALYSIS

Think of the unconstrained model as a **performance ceiling**

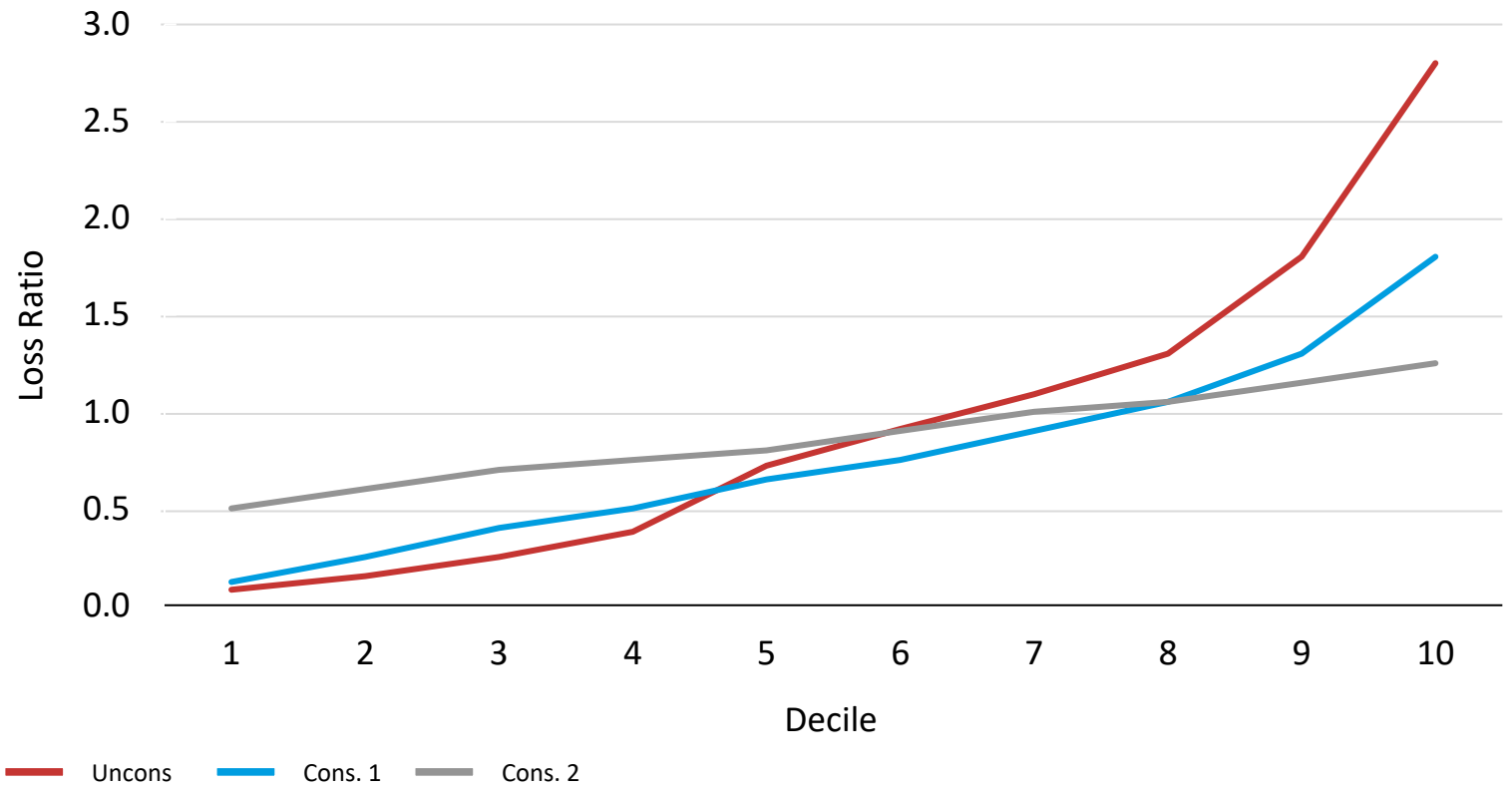
- Benchmark other models against this ceiling
- This represents the maximum value offered by your data

What will the gap reveal?

- Which constraints cost the most predictive power?
- Are any constraints correlated with each other?
- Which constraints are required?
- How much of the gap is unavoidable?

STRATEGIC VALUE OF MODELS

Example lift chart with 3 models | Gap Analysis – Red vs. Blue | Gap Analysis – Red vs. Gray



Insights

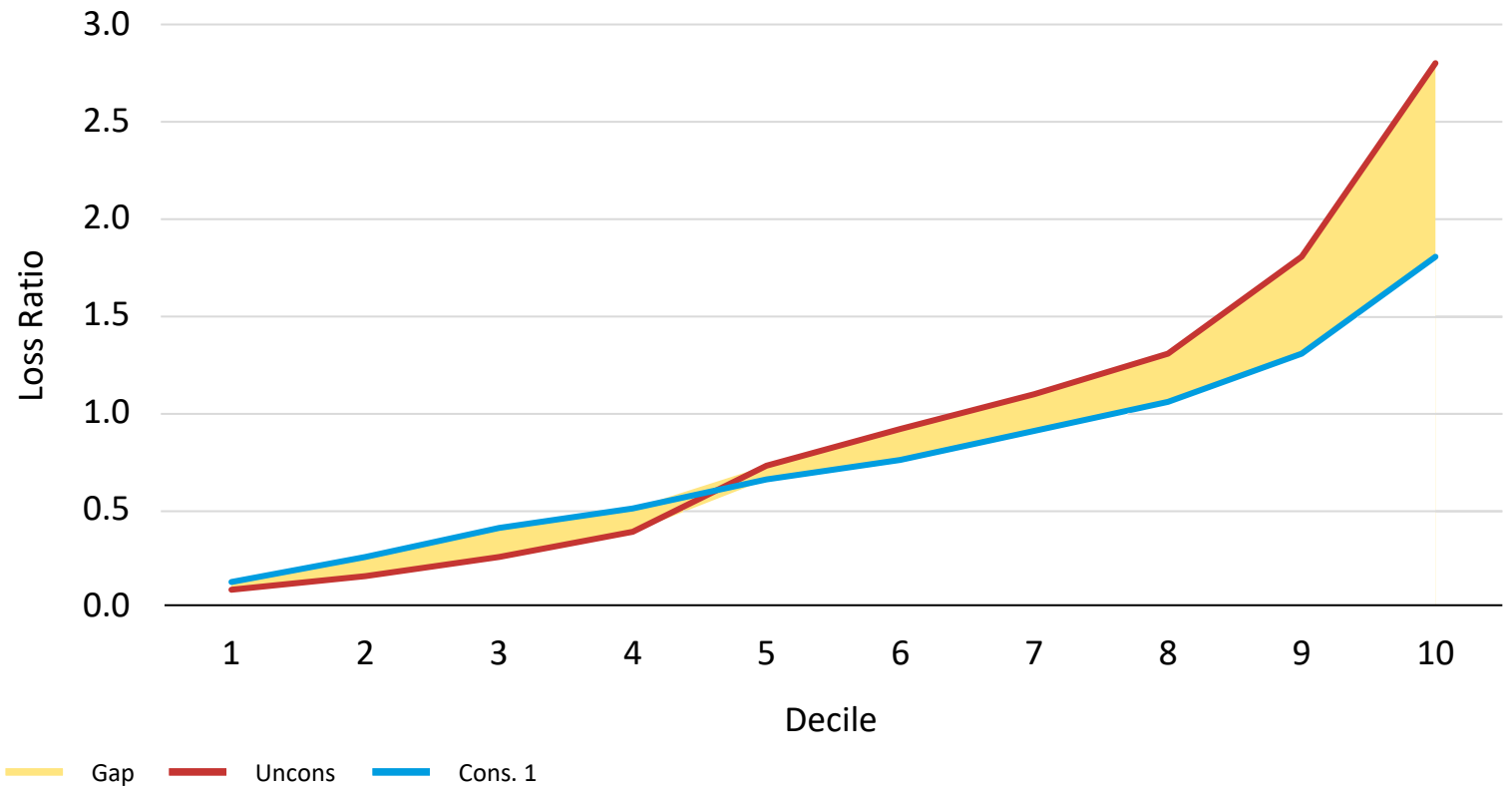
- Good models appear as increasing reading the chart left to right
- Models

Red	Unconstrained (best)
Blue	Moderately constrained model
Gray	Highly constrained model (worst)

All models should aspire to be the Red model – it has the highest predictive power of all models considered

STRATEGIC VALUE OF MODELS

Example lift chart with 3 models | **Gap Analysis – Red vs. Blue** | Gap Analysis – Red vs. Gray



Insights

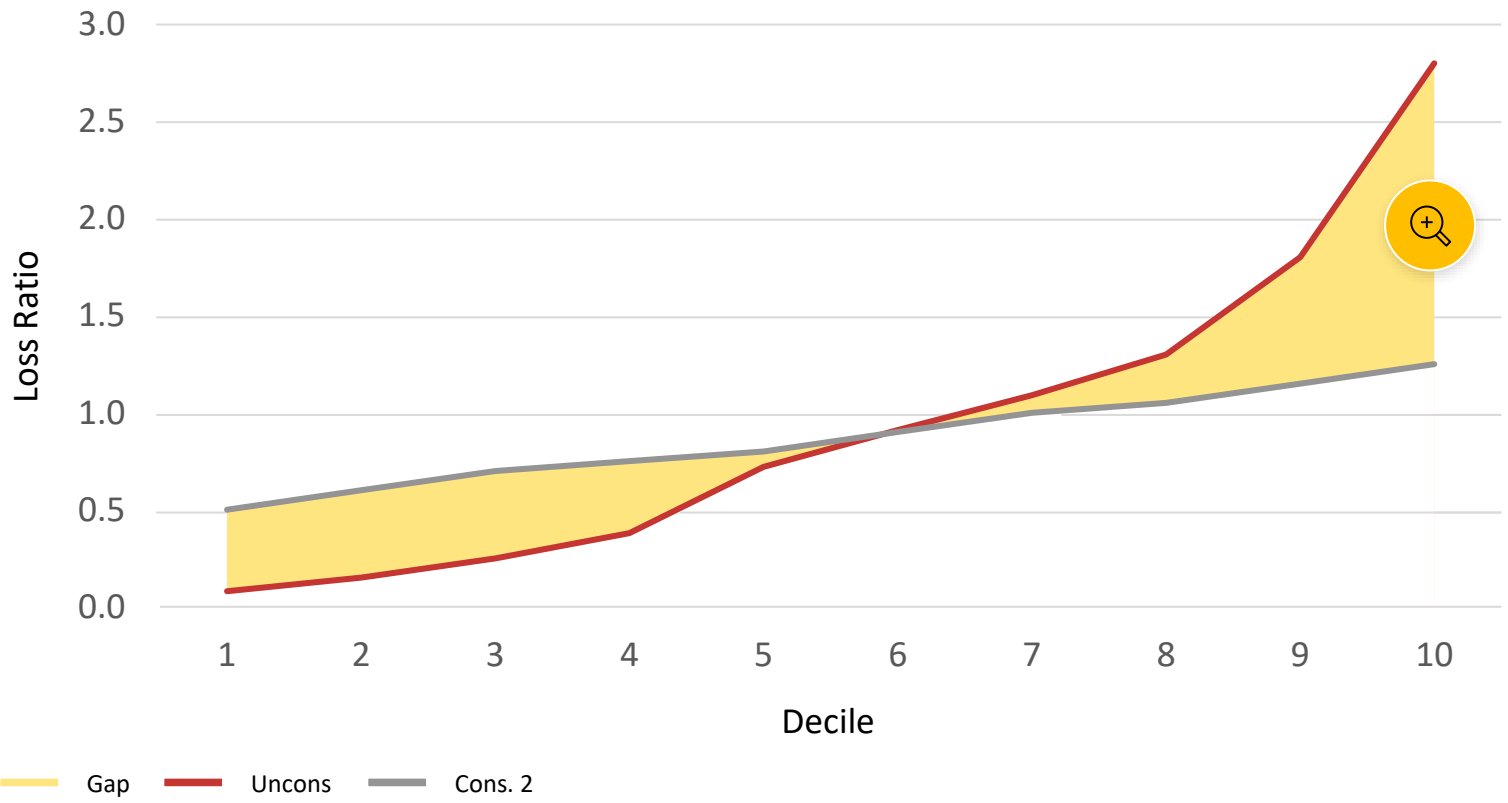
- The **yellow area "gap" is the predictive power lost from adding constraints**
- The model has lost segmentation power through the middle and especially in the right side of the lift chart

How much was lost?

57%	Lift	Compares endpoints
24%	Area	Compares areas under curve

STRATEGIC VALUE OF MODELS

Example lift chart with 3 models | Gap Analysis – Red vs. Blue | **Gap Analysis – Red vs. Gray**



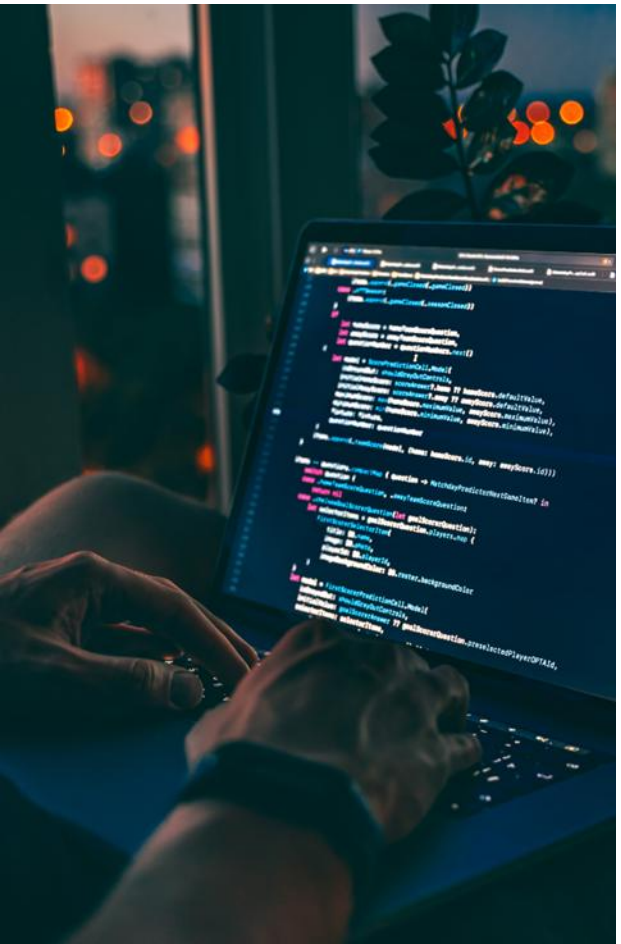
Insights

How much was lost?

93% Lift | **43%** Area

- Predictive Power
- Fairness
- Segmentation
- Insights
- Value
- Profit

CONSTRAINTS AND MODEL FLEXIBILITY



Customer Experience

- Identify events leading to churn
- Loyalty and rewards programs

Risk Identification and Mitigation

- Risk segmentation
- Risk management
- Underwriting guidelines

Claims Management

- Case reserves
- Triage
- Special investigation
- Fraud detection

Operational Efficiency

- Call center/staffing resource allocation
- Data source evaluation

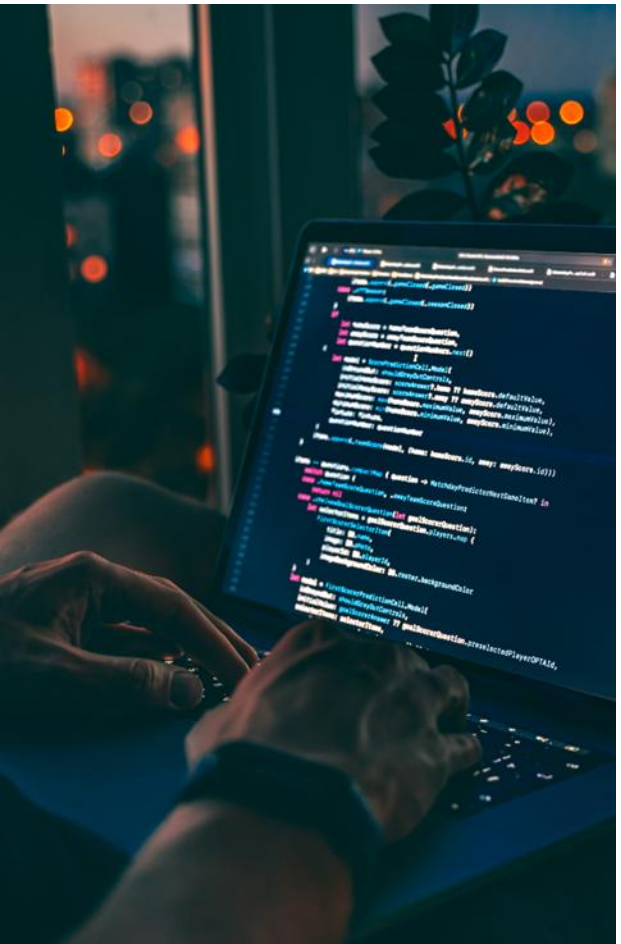
Reserving

- Claim-level reserving
- Business mix optimization

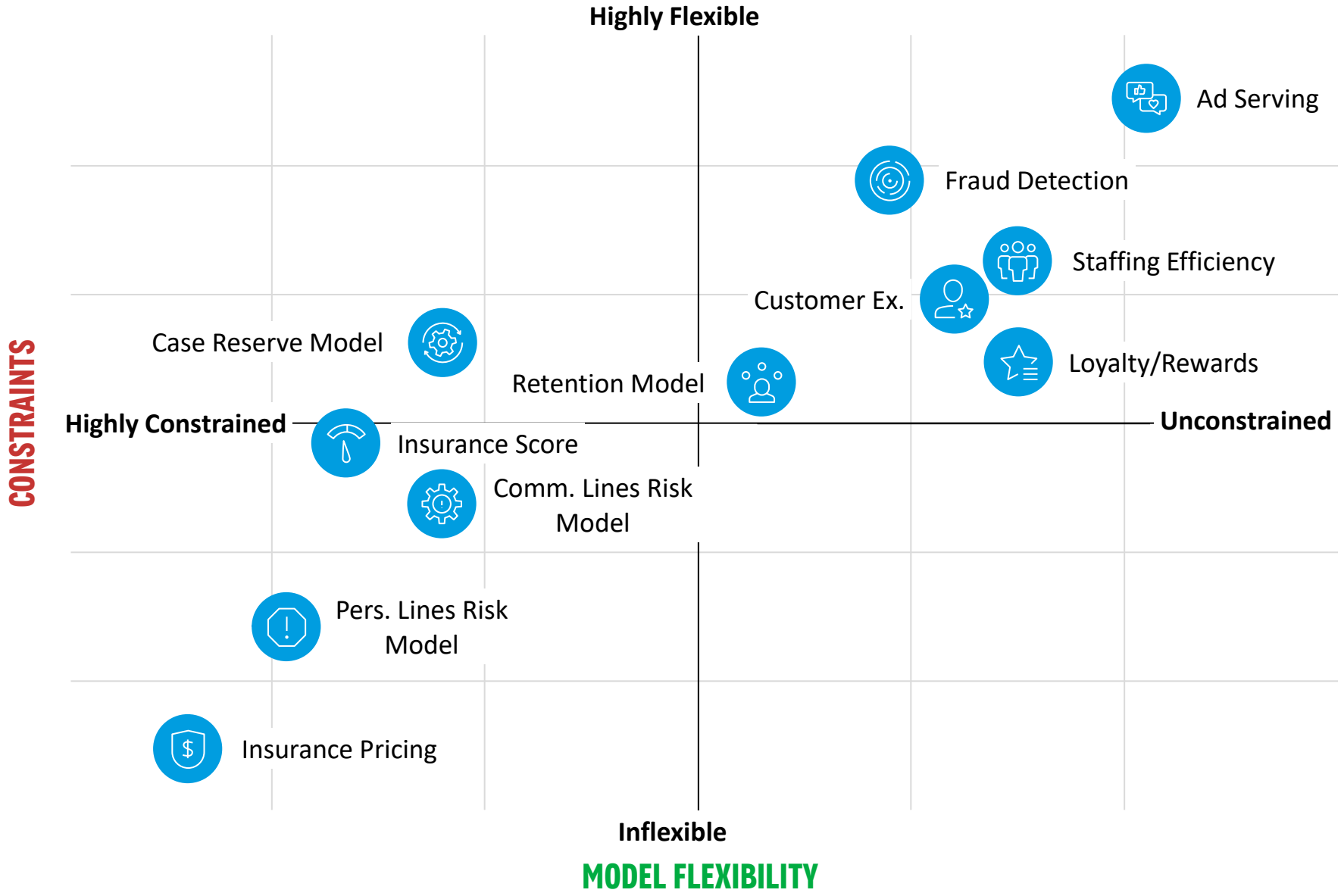
Marketing

- Serve ads to preferred business
- Evaluate profitable partnerships

CONSTRAINTS AND MODEL FLEXIBILITY

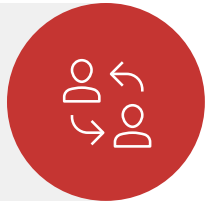


Where does your next project fall on our chart?



EXAMPLES OF ROADBLOCKS

Examples of different models which led to adverse results – they likely did not consider all appropriate constraints



RECRUITING

Models that process job applications and resumes for hiring trained on historical data will incorporate current and historical hiring biases

REAL-TIME DEMAND-INFLUENCED PRICING

This kind of price setting may increase prices during times when people are very vulnerable (natural disasters, mass shootings), which may increase profits but could clash with cultures and values

RECOMMENDATION ALGORITHMS

Algorithms that recommend content to prioritize engagement – this can show controversial, graphic, or inappropriate content – which can create a feedback loop that leads to more similar content

INCENTIVE PROGRAMS

Tying pay, bonuses, or other compensation to applications, accounts, or policies closed may cause staff to engage in activities to maximize that metric to detriment of customers

AI GOVERNANCE (1 OF 2)

1

Accountability

- Consider an overarching AI Governance Committee with cross-functional representation
- Establish clear ethical standards for AI development
- Traceability and Auditing for AI models and tools

2

Transparency

- Document data sources used for AI training and inferencing
- Assess and prioritize risks based on the AI system life cycle
- Consider transparency and interpretability especially in high-stake decisions

3

Privacy

- Minimize collection and use of sensitive data
- Where applicable, implement informed consent
- Proactively track emerging regulations

AI GOVERNANCE (2 OF 2)

4 **Fairness**

- Test for and address bias across the lifecycle
- Verify AI-supported decisions do not lead to discriminatory outcomes
- Create a process to immediately address instances of bias or unfair outcomes

5 **Security & Safety**

- Ensure AI infrastructure is secure
- Determine safeguards to prevent or contain worst-case scenarios
- Ensure vendors and partners align to internal standards
- Create continuous monitoring

6 **Third-Party Risk**

- Opaque model training practices with risks of bias, IP violations, and regulatory exposure
- Unmonitored model drift and unclear model versioning
- Embedded AI in off-the-shelf tools
- Inadequate security protocols

GETTING STARTED WITH UNCONSTRAINED MODELS

Data Requirements

- Need data, lots of it.
  - Internal
    - Loss and exposure data, transactional
  - Publicly available
    - Census, FEMA, Economic
  - Licensed
    - Satellite imagery, risk scores, market basket

Technology and Tools

- Predictive modeling software
  - Python/R, Akur8, Alteryx, SQL
- Dashboarding and analytics tools
  - Python/R, Power BI, Tableau, Excel

Team and Talent Collaboration

- Multidisciplinary team
  - Data scientists
  - Data engineers
  - Developers
  - Actuaries

Iterative Development and Monitoring

- Predictive model process is iterative
  - Data -> Train -> Evaluate, Repeat
- Monitor data for
  - Trends and shifts
  - Quality
  - Opportunities

SUMMARY SLIDE

Q&A W/ PAT

QUALIFICATIONS, ASSUMPTIONS, AND LIMITING CONDITIONS

This report is for the exclusive use of the Oliver Wyman client named herein. This report is not intended for general circulation or publication, nor is it to be reproduced, quoted, or distributed for any purpose without the prior written permission of Oliver Wyman. There are no third-party beneficiaries with respect to this report, and Oliver Wyman does not accept any liability to any third party.

Information furnished by others, upon which all or portions of this report are based, is believed to be reliable but has not been independently verified, unless otherwise expressly indicated. Public information and industry and statistical data are from sources we deem to be reliable; however, we make no representation as to the accuracy or completeness of such information. The findings contained in this report may contain predictions based on current data and historical trends. Any such predictions are subject to inherent risks and uncertainties. Oliver Wyman accepts no responsibility for actual results or future events.

The opinions expressed in this report are valid only for the purpose stated herein and as of the date of this report. No obligation is assumed to revise this report to reflect changes, events, or conditions, which occur subsequent to the date hereof.

All decisions in connection with the implementation or use of advice or recommendations contained in this report are the sole responsibility of the client. This report does not represent investment advice nor does it provide an opinion regarding the fairness of any transaction to any and all parties. In addition, this report does not represent legal, medical, accounting, safety, or other specialized advice. For any such advice, Oliver Wyman recommends seeking and obtaining advice from a qualified professional.

CASE STUDY

Staffing Optimization Model

GOAL

- Provide staffing numbers needed for global call centers to handle claims intake for a non-standard auto insurer
- Automated – move away from a manual process
- Responsive to changes in business

Considerations

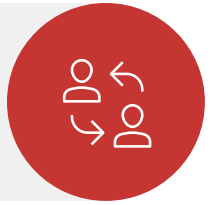
Work from home vs. in-office:	There is a max number of people that should be in a building
Training:	Workers need training to do their job well
Cost:	Cannot hire an unlimited amount of people
HR:	Cannot/should not fire people suddenly
Scale with company growth:	More claims are expected as policies are written

An unconstrained model in this context serves multiple purposes

Benchmarking	**Sensitivity:**	**Value:**
What is possible without limiting the inputs/outputs	Which inputs cause large changes in the outputs	Which combination of inputs maximizes financial metric

MODELS THAT WENT WRONG

Examples of different models which led to adverse results – they likely did not consider all appropriate constraints



RECRUITING

Models that process job applications and resumes for hiring trained on historical data will incorporate current and historical hiring biases

REAL-TIME DEMAND-INFLUENCED PRICING

This kind of price setting may increase prices during times when people are very vulnerable (natural disasters, mass shootings), which may increase profits but could clash with cultures and values

RECOMMENDATION ALGORITHMS

Algorithms that recommend content prioritize engagement – this can show controversial, graphic, or inappropriate content – which can create a feedback loop that leads to more similar content

INCENTIVE PROGRAMS

Tying pay, bonuses, or other compensation to applications, accounts, or policies closed may cause staff to engage in activities to maximize that metric to detriment of customers