Analytics in General Insurance

28 October 2021

Predictive Modelling: Way forward - GLM with Actuarial Judgment or Machine Learning Models

Institute of Actuaries of India

Introductions

Neil Chapman

Senior Director P&C UK,
Global leader for Pricing Product Claim and Underwriting

Kylie Chen

Director P&C APAC,
Pricing Product Claim and Underwriting leader for APAC

Sipika Tandon

Associate Director P&C India

Agenda

Institute of Actuaries of India

Agenda

- **Context of machine learning in pricing**
- **Introduction to machine learning and comparison to GLMs**
  - **Tree-based methods**
    - **Decision trees**
    - **Random forests**
    - **Gradient boosting machines**
  - **Regression-based methods**
    - **Penalized regression**
- **Pricing applications**
- **How are insurers using machine learning?**
- **Adoption and use of machine learning in India and APAC**
- **Summary and Q&A**

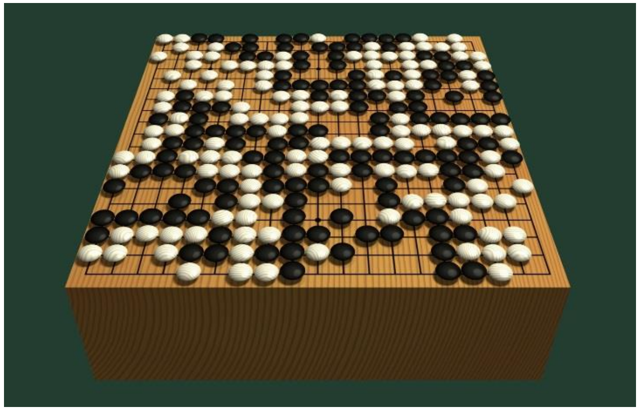
Objective: to understand the advantages and disadvantages of machine learning, and how these could be used to enhance predictive modelling alongside actuarial judgement

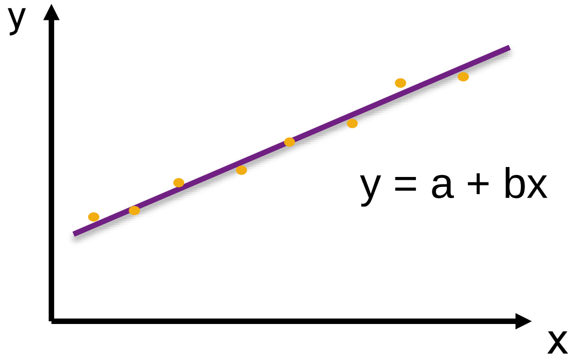
[redacted]

Context of machine learning in pricing

Who's interested in what?

Institute of Actuaries of India



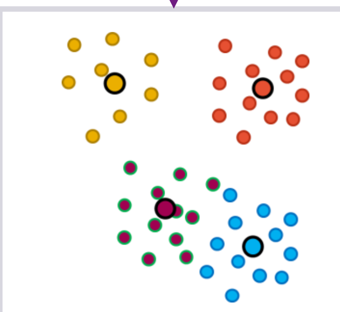


Institute of Actuaries of India

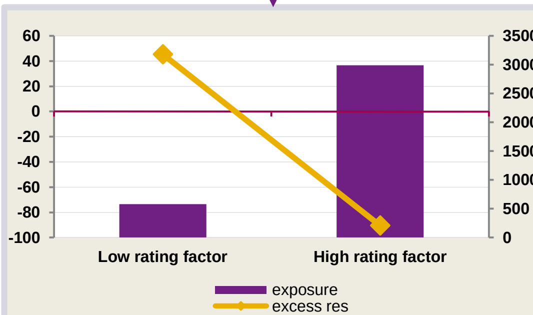
Applications of machine learning in the insurance sector

Marketing and Distribution → Customer management → Underwriting and risk management → Pricing → Claims management → Asset management

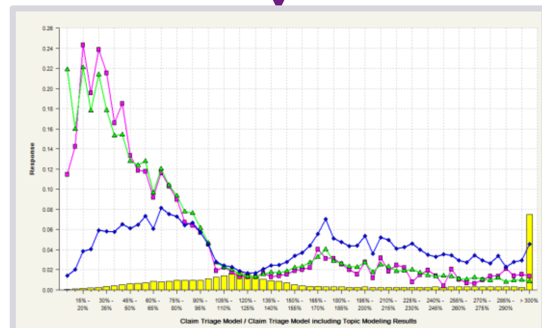
- Product development (UBI)
- Retention segmentation and program design
- Agent/Broker performance evaluation
- Fraud detection



P&C to Life cross-selling



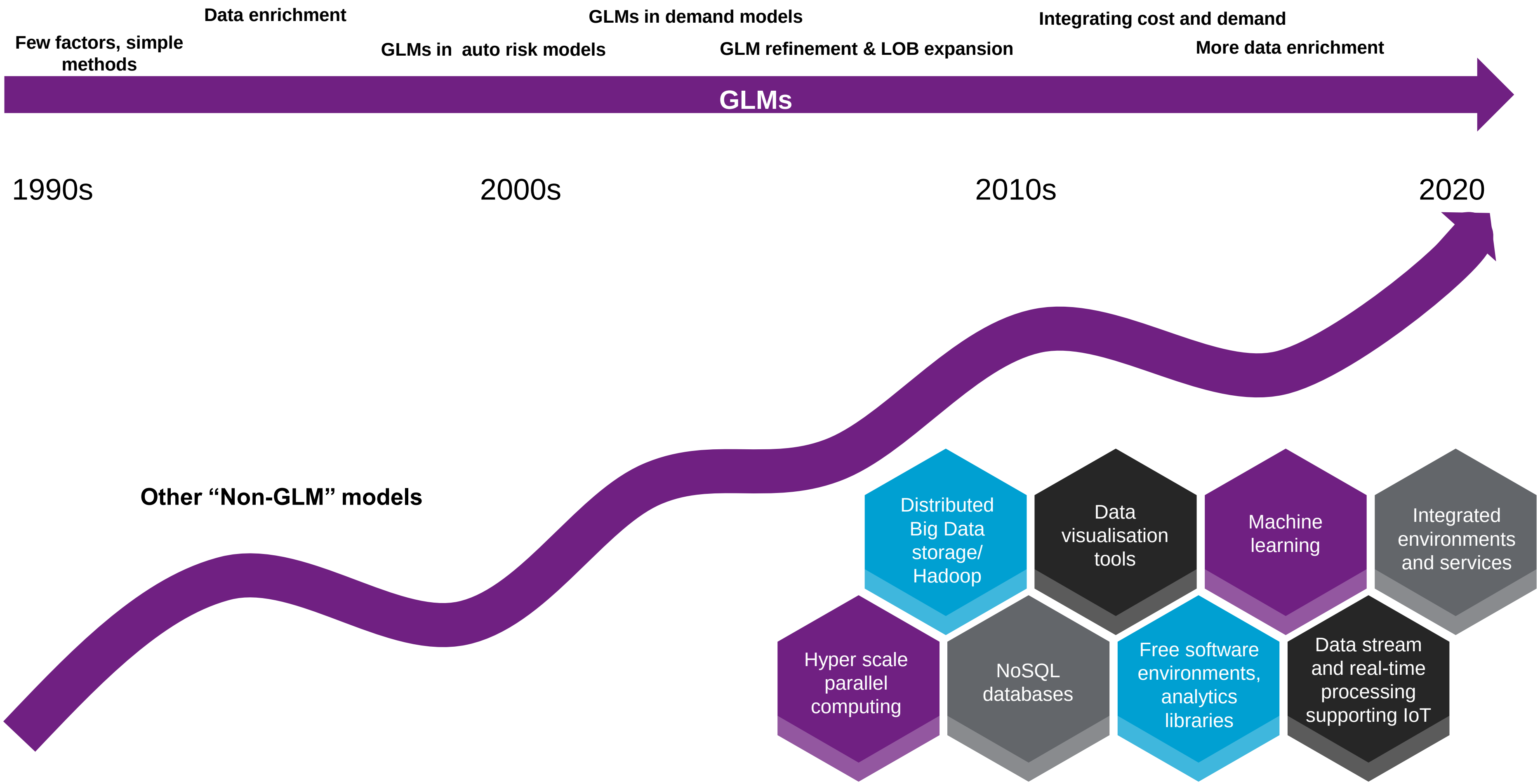
Web-scraping and feature design in commercial lines



Topic modeling and large loss modeling

This is not new….

Institute of Actuaries of India



What are these machine learning methods?

Ensembles

Classification Trees

"Earth"

Regression Trees

Gradient Boosting Machines

K-nearest Neighbors

Elastic Net

Neural Networks

Naïve Bayes

Random Forests

K-Means Clustering

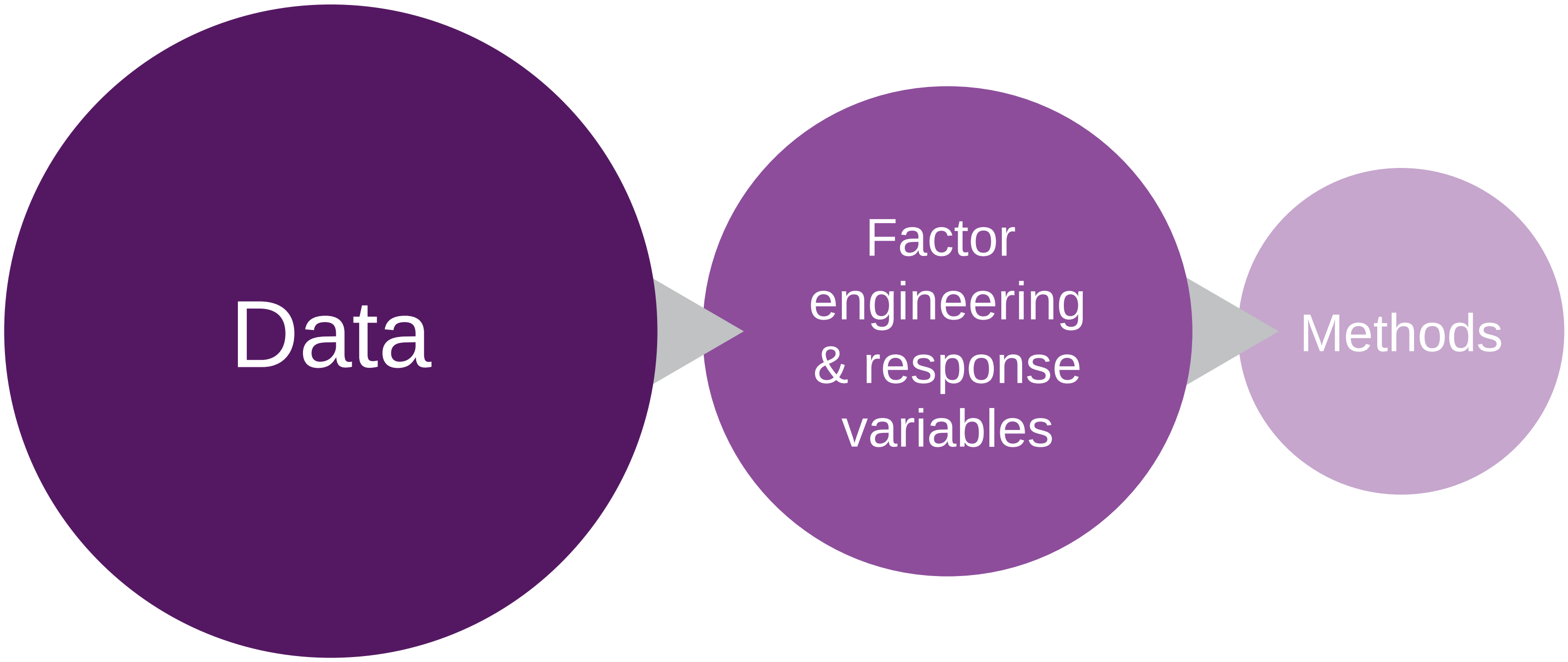
Principal Components Analysis

Lasso

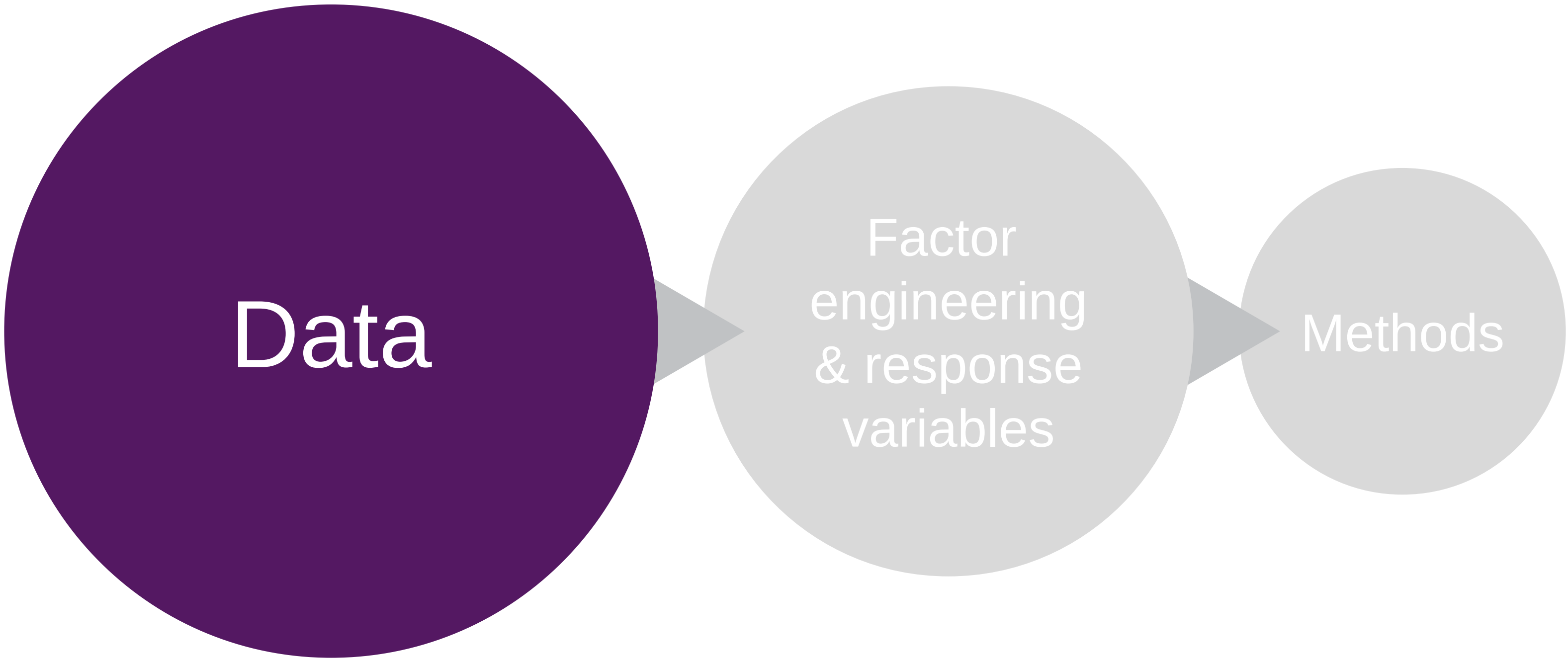
Support Vector Machines

Ridge Regression

Is it really all about the method?



Is it really all about the method?

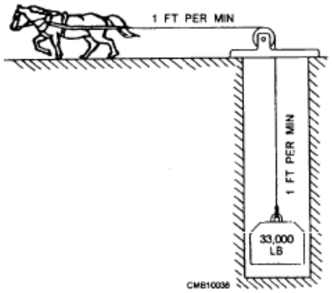


Is it really all about the method?

Data

Physical facticity
E.g., height, length, weight

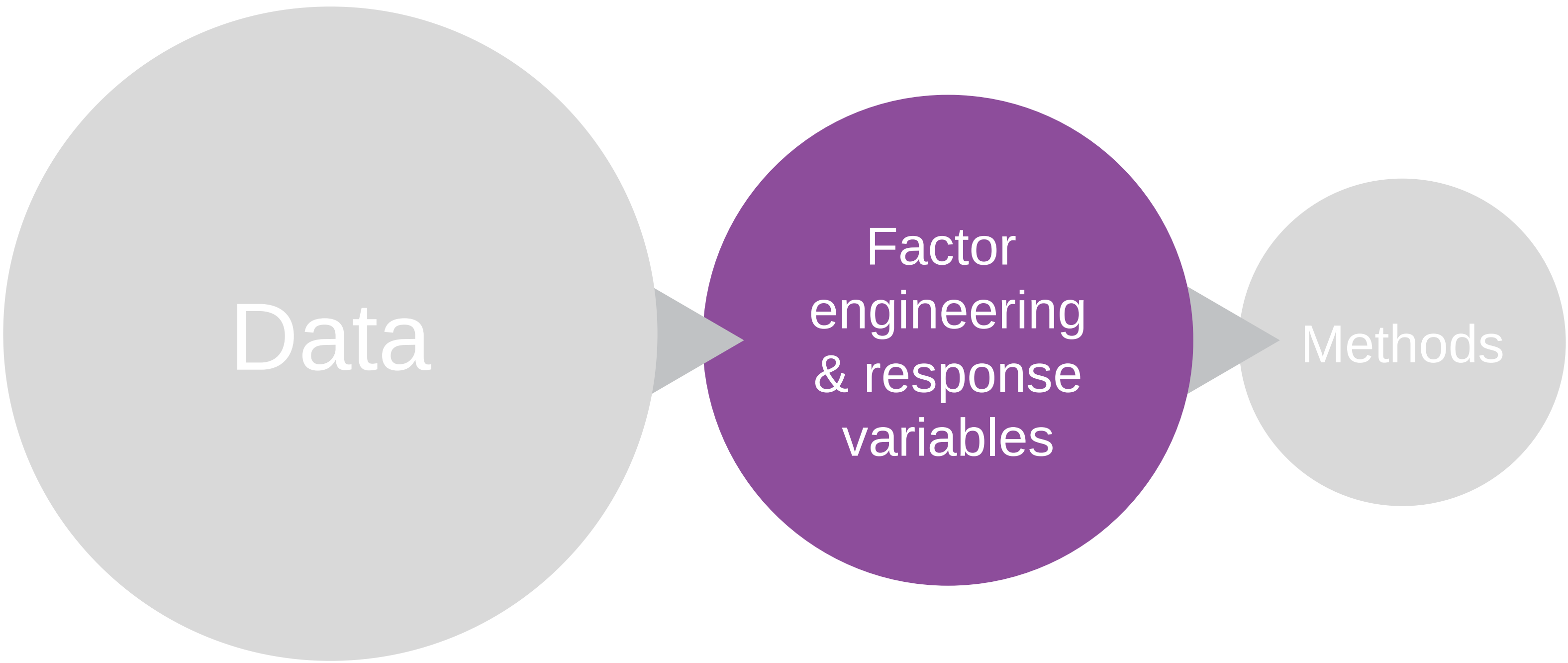
Mechanical nature
E.g., engine size, fuel type



Qualitative descriptors
E.g., body type, model range

Performance
E.g., maximum speed, torque, BHP

Is it really all about the method?

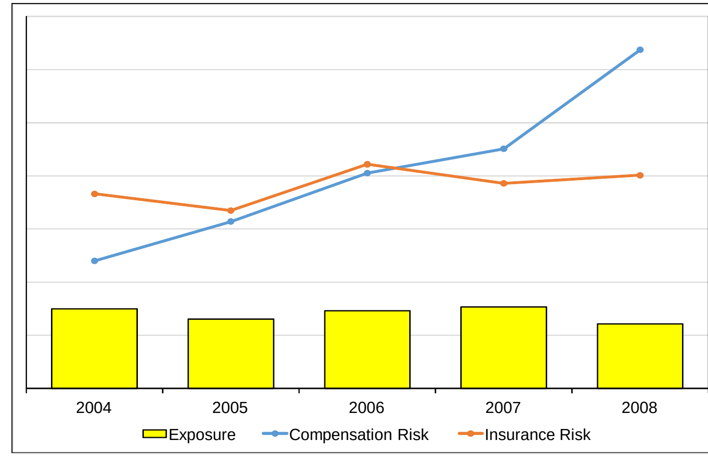


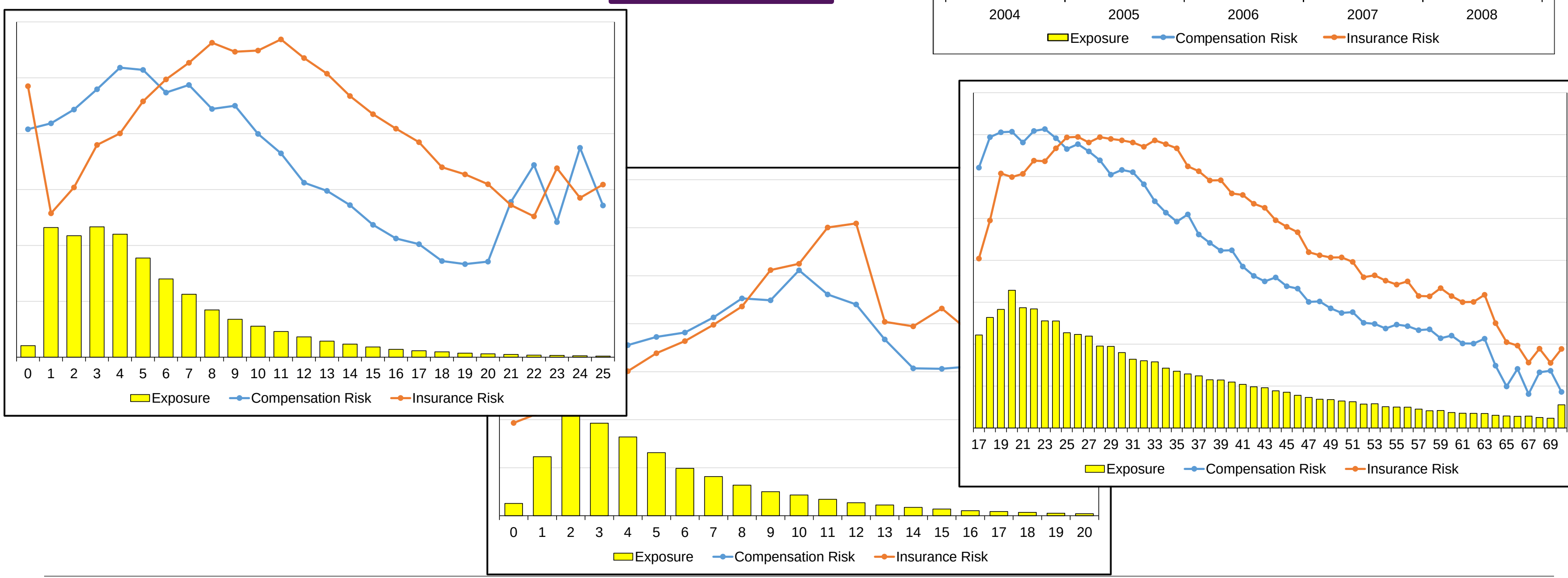
Institute of Actuaries of India

Is it really all about the method?

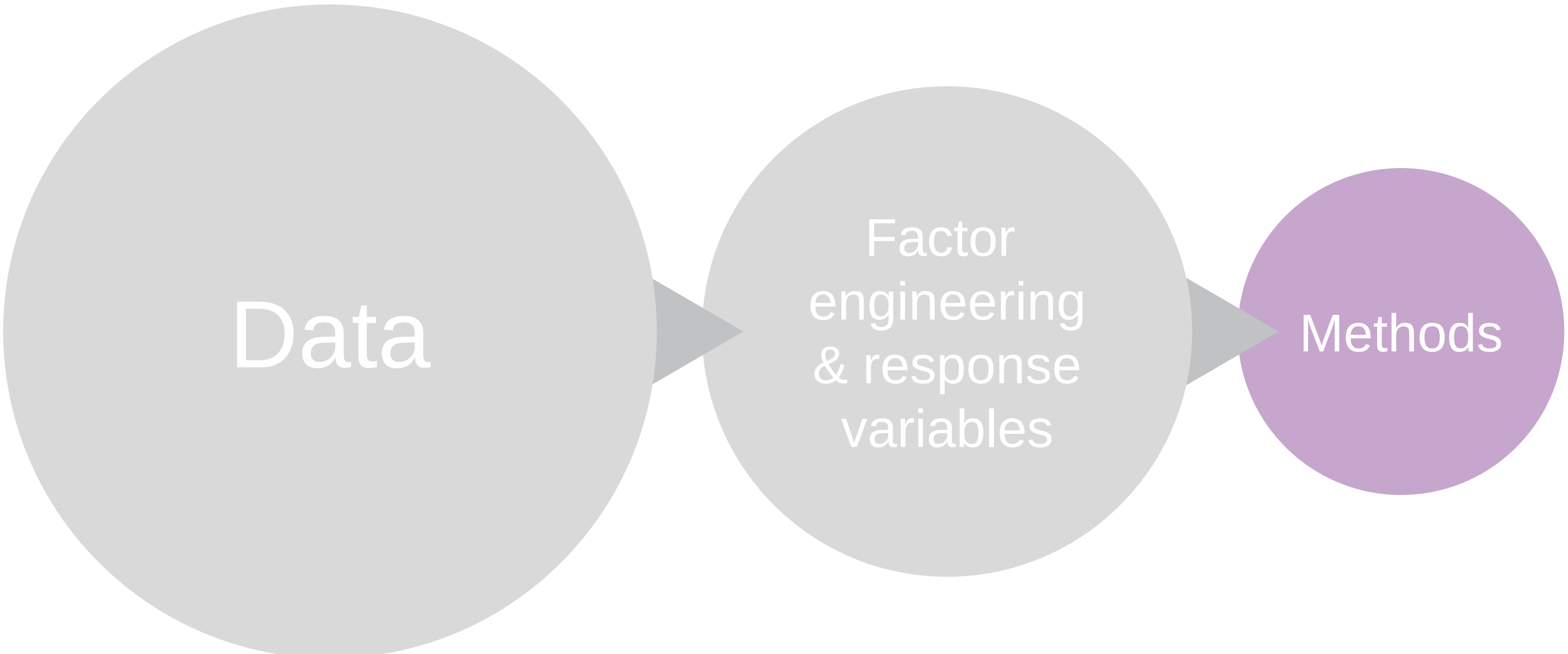
Response selection

Bodily Injury → Simple guessed algorithm → "Insurance risk" / "Compensation risk"





Is it really all about the method?

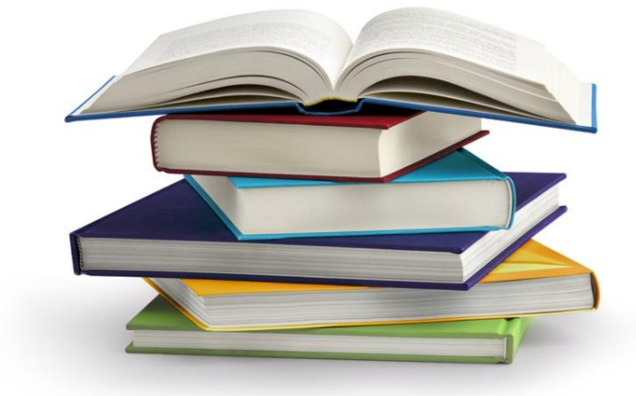


How do you know if a method works?

AIC

Gini

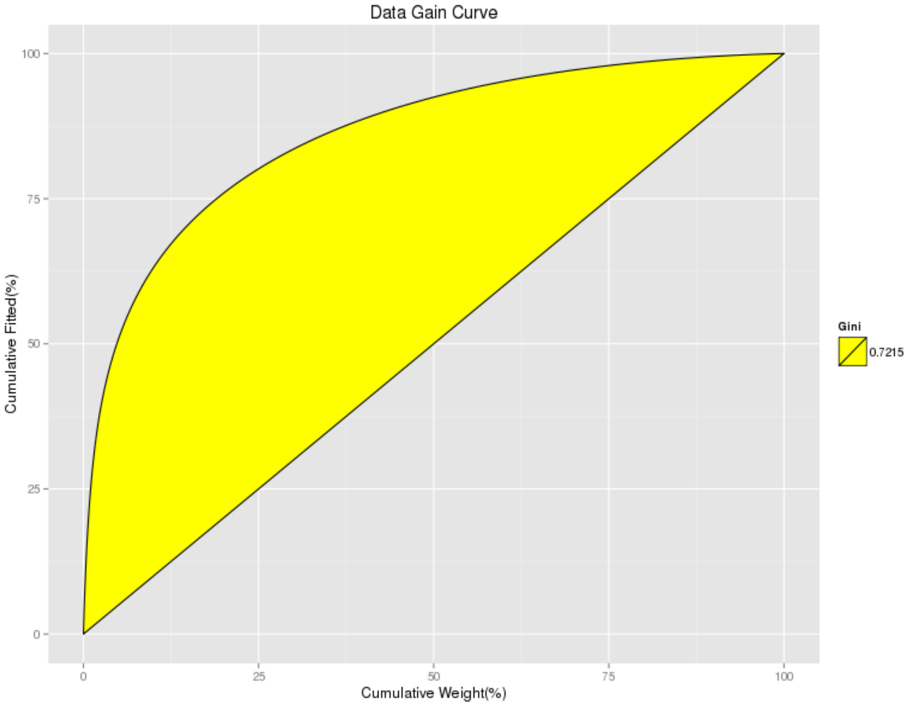
RMSE



MAE

Log
loss

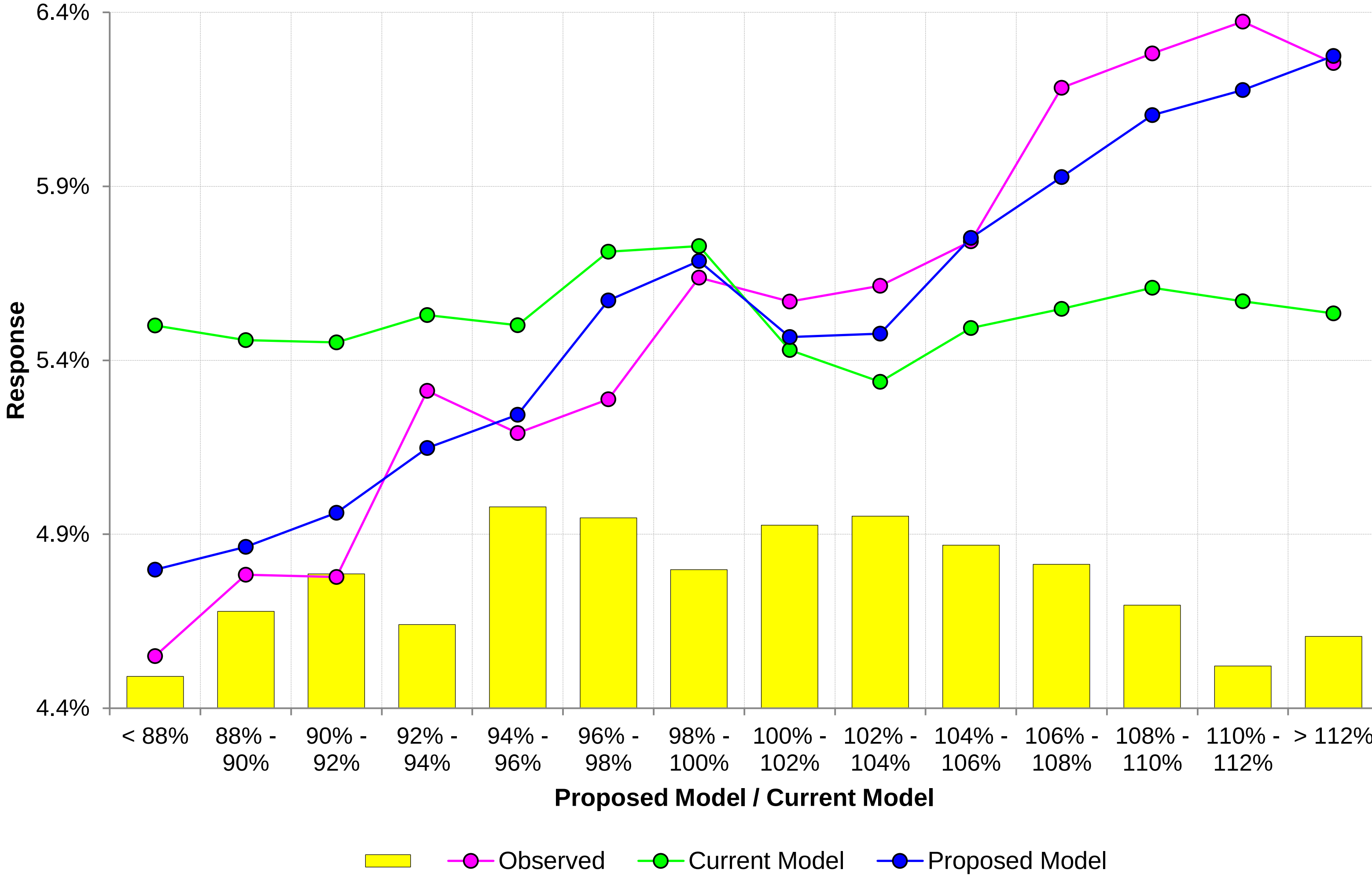
How do you measure value?



- Rank hold out observations by their **fitted values** (high to low)
- **Plot cumulative response** by cumulative exposure
- A **better model** will explain a **higher proportion of the response** with a **lower proportion of exposure**
- …and will give a **higher Gini coefficient** (yellow area)

Double lift chart

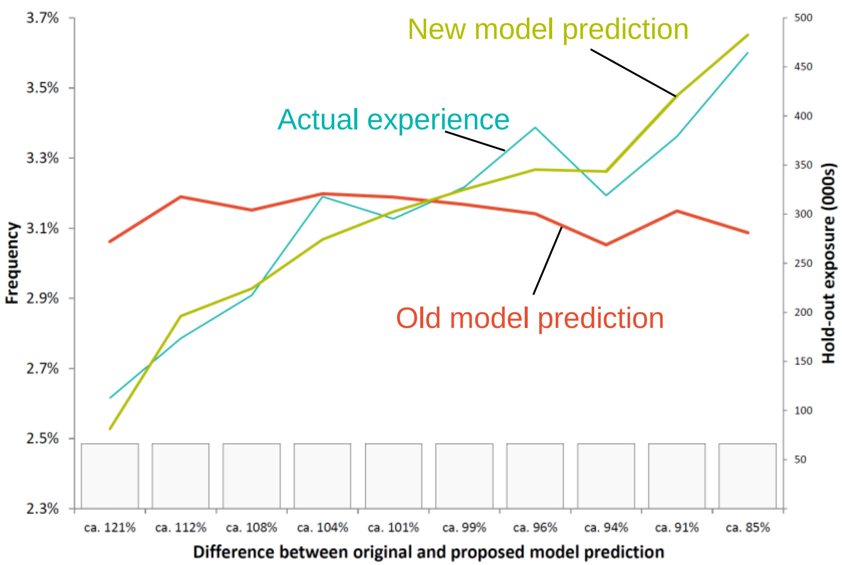
Institute of Actuaries of India



Financial value estimate

Institute of Actuaries of India

- Errors in insurance pricing are not symmetrical
- Financial benefit can be estimated
- Consider ***actual experience*** in out of sample data for each percentile of old vs new model fitted values
- Estimate financial benefit that would have been attained
  - given an assumed elasticity
  - given business rules such as an assumed cap/floor approach

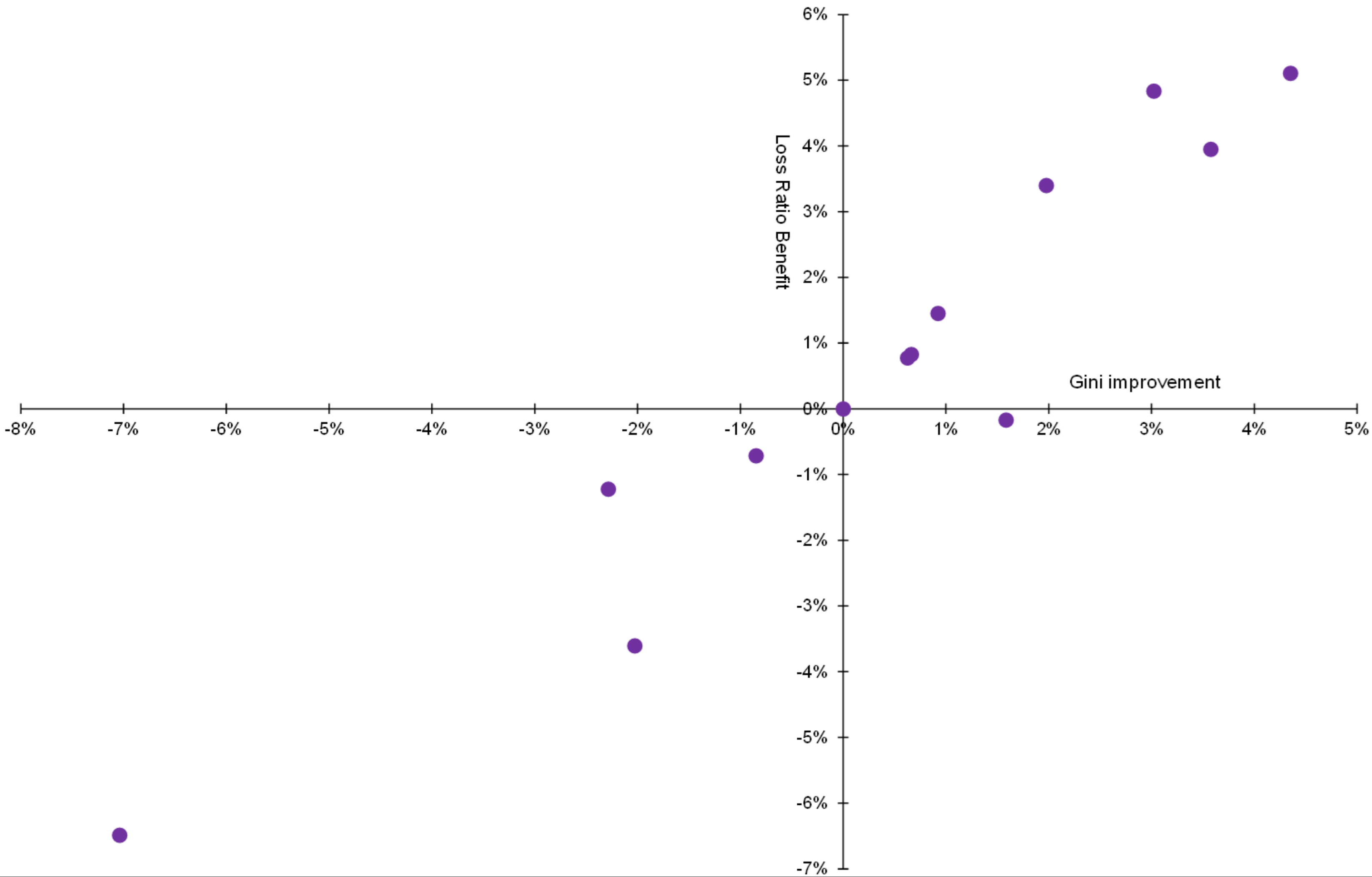


Simple formula

Old/New	New premium	Expected volume	Actual claims	Increased profit
121%	P_1	V_1	C_1	X_1
…	P_2	V_2	C_2	X_2
…	…	…	…	…
…	P_{99}	V_{99}	C_{99}	X_{99}
85%	P_{100}	V_{100}	C_{100}	X_{100}
			Value created	$ X

Financial value vs Gini

Institute of Actuaries of India



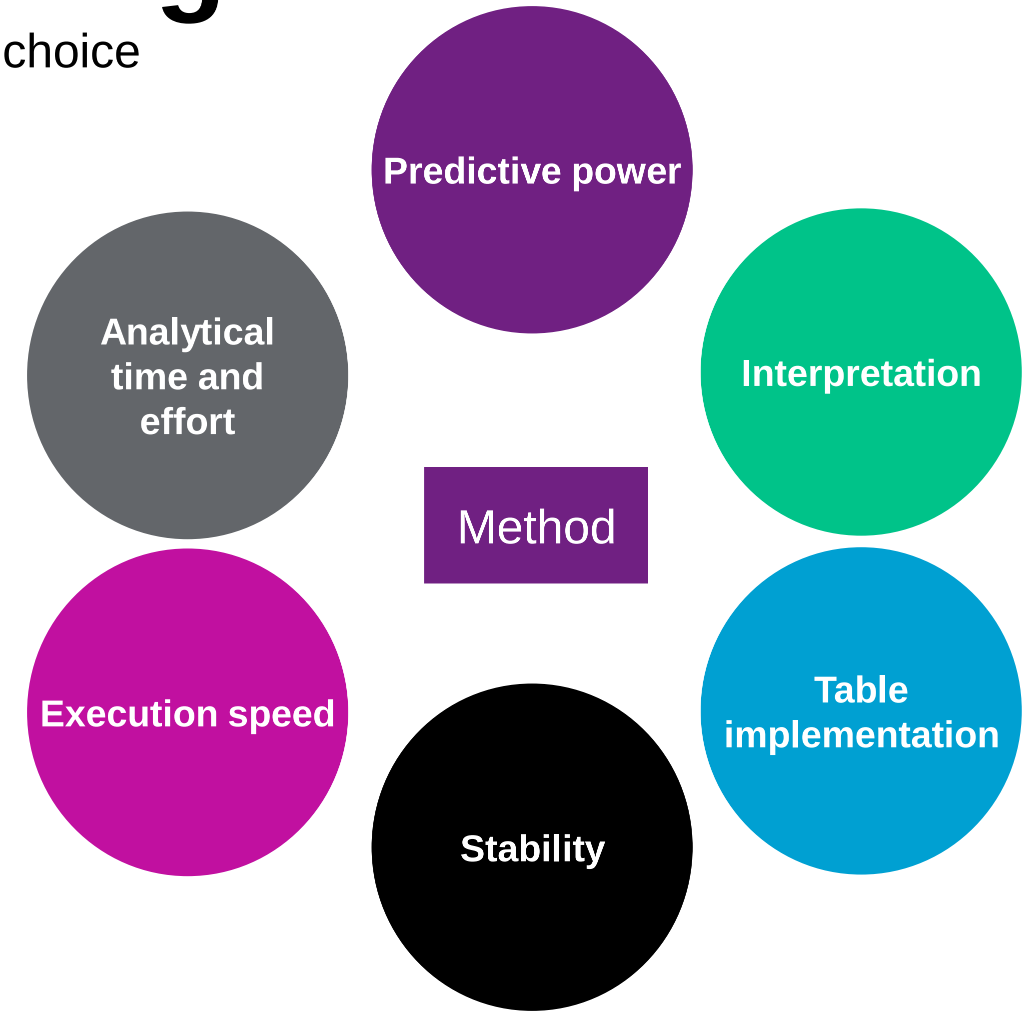
Institute of Actuaries of India

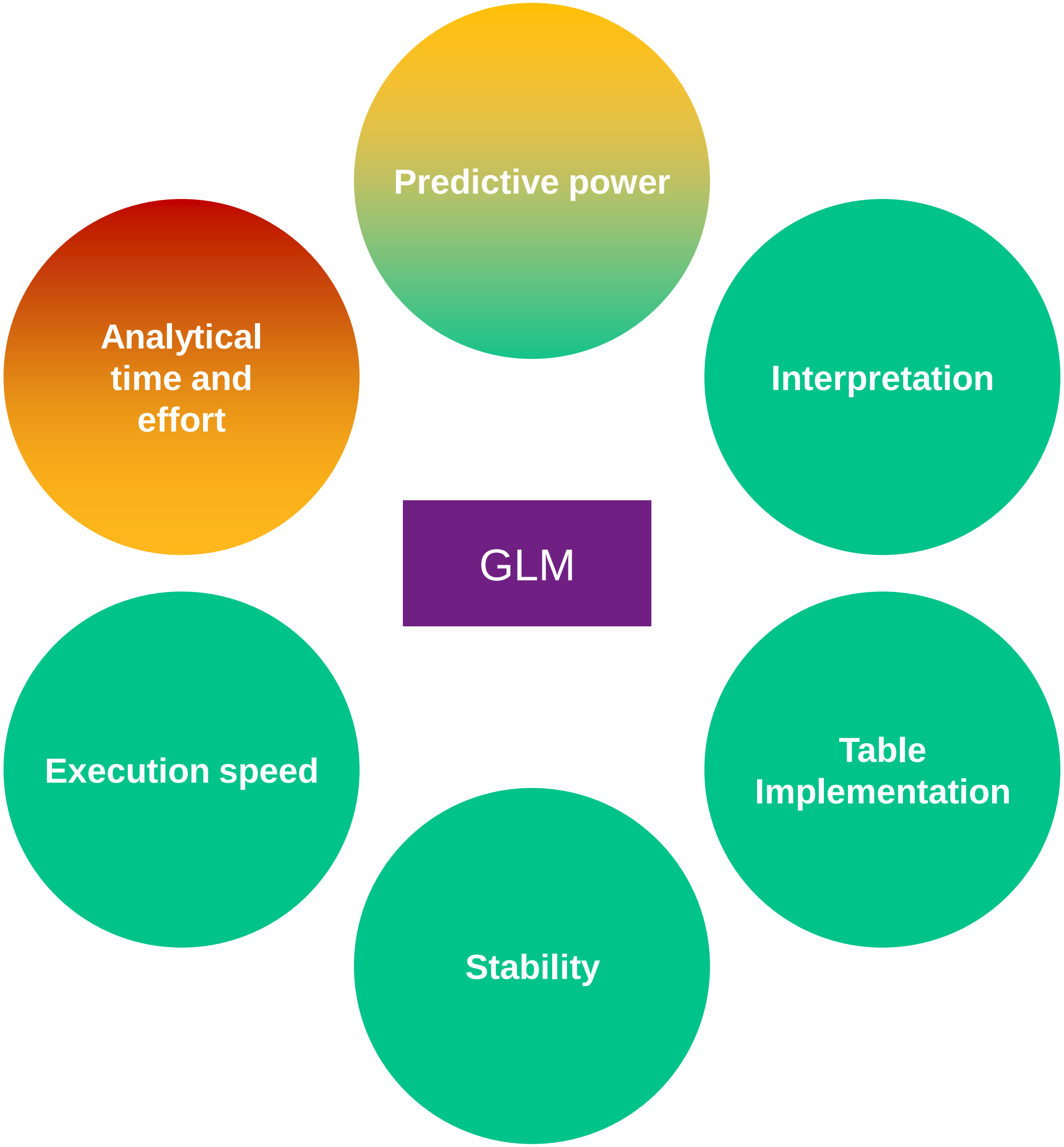
Is there more to it…?

Predictive power

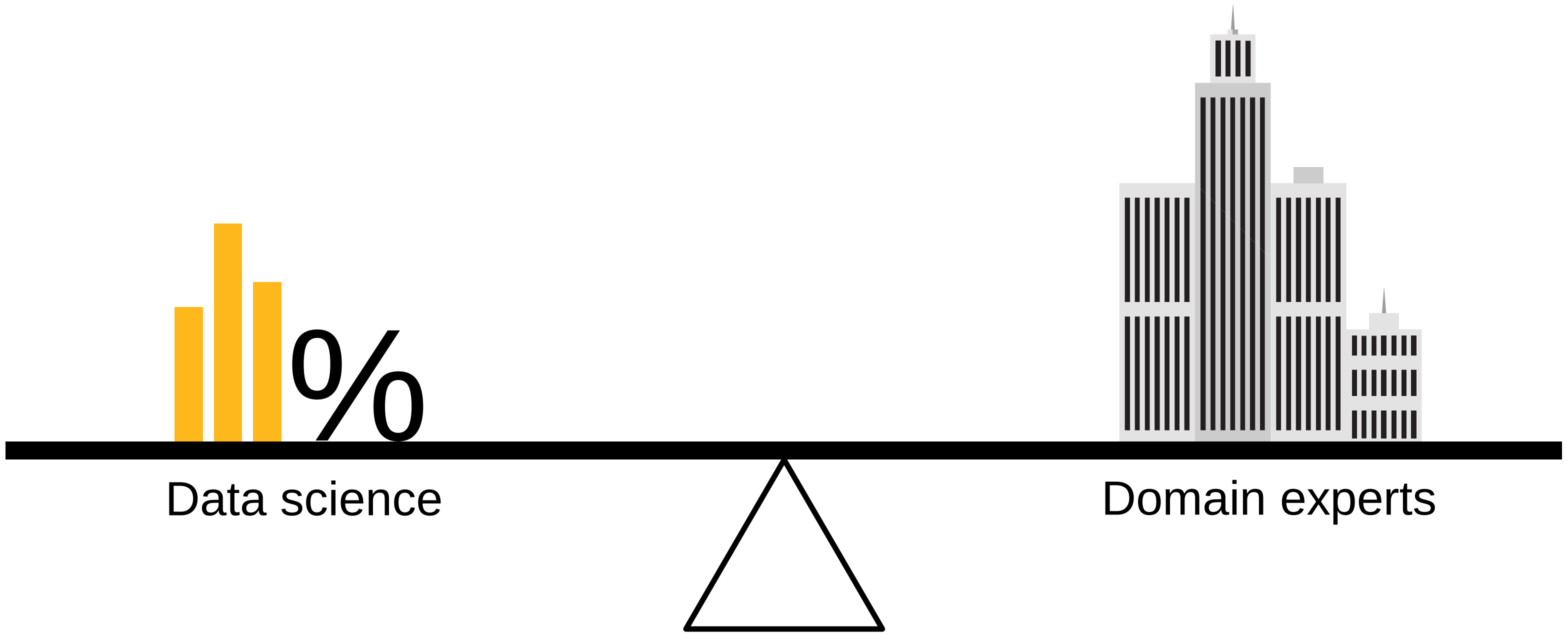
Choosing a method

Dimensions of choice

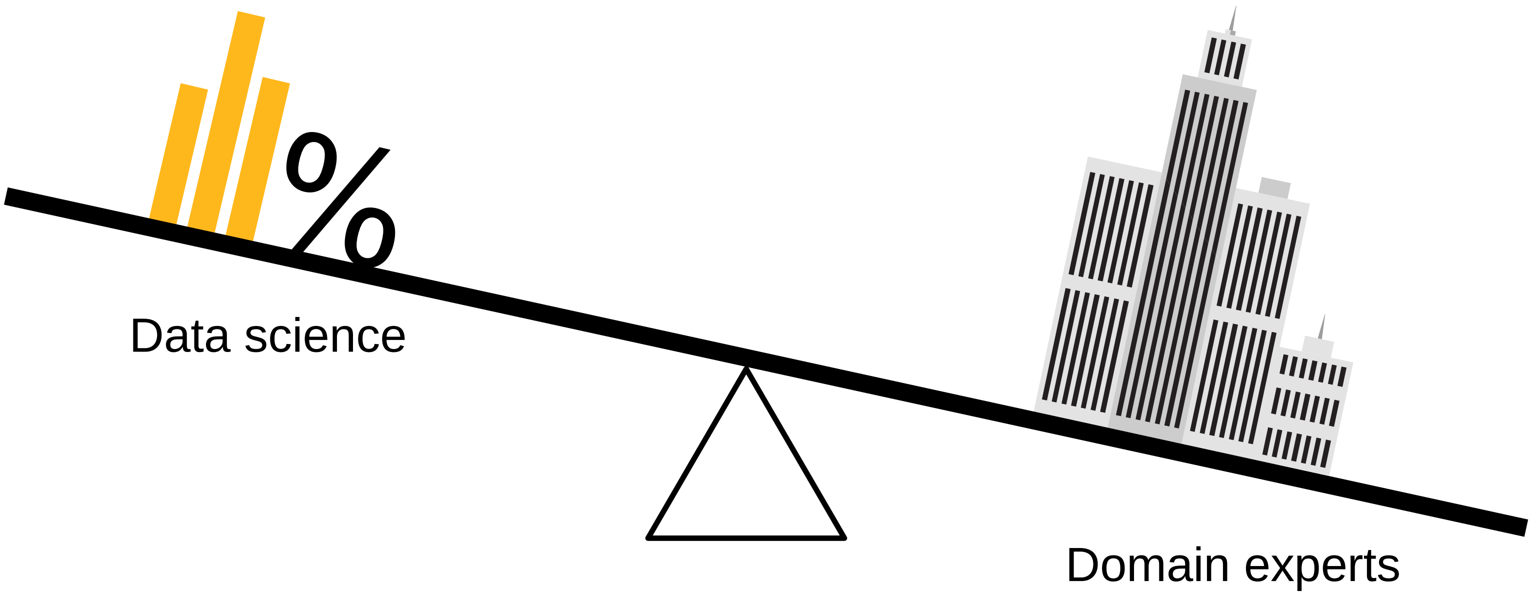



Predictive power
Analytical time and effort
Interpretation
GLM
Execution speed
Table Implementation
Stability

What do you use where?



It's domain expertise that helps decide



Tree-based methods

Institute of Actuaries of India

Some machine learning methods

Ensembles	Classification Trees	"Earth"	Regression Trees	Gradient Boosting Machines
K-nearest Neighbors	Elastic Net	Neural Networks	Naïve Bayes	Random Forests
K-Means Clustering	Principal Components Analysis	Lasso	Support Vector Machines	Ridge Regression

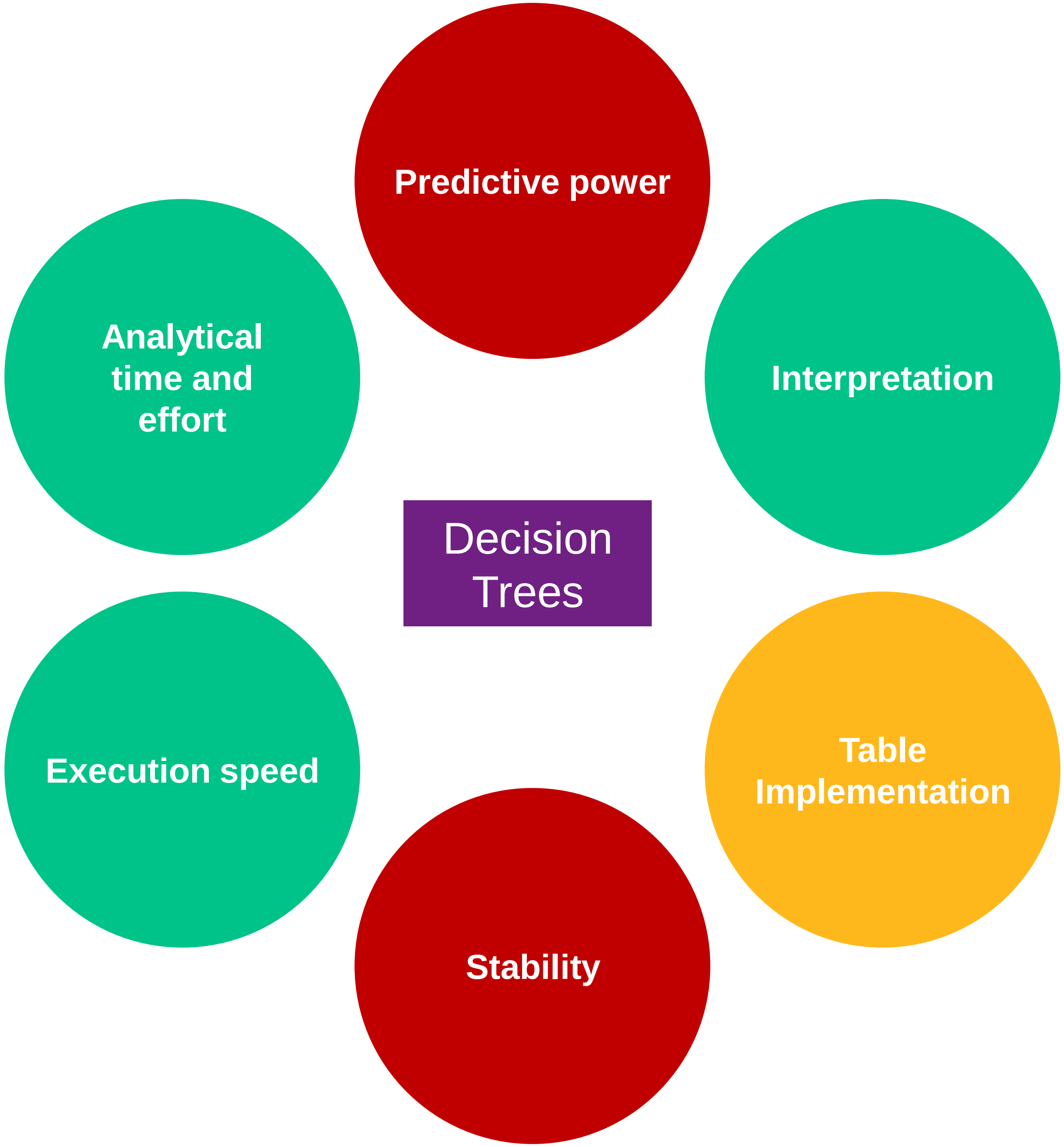
Institute of Actuaries of India

Focus on Trees

Classification Trees

Regression Trees

Institute of Actuaries of India



Some machine learning methods

Ensembles

Classification Trees

"Earth"

Regression Trees

Gradient Boosting Machines

K-nearest Neighbors

Elastic Net

Neural Networks

Naïve Bayes

Random Forests

K-Means Clustering

Principal Components Analysis

Lasso

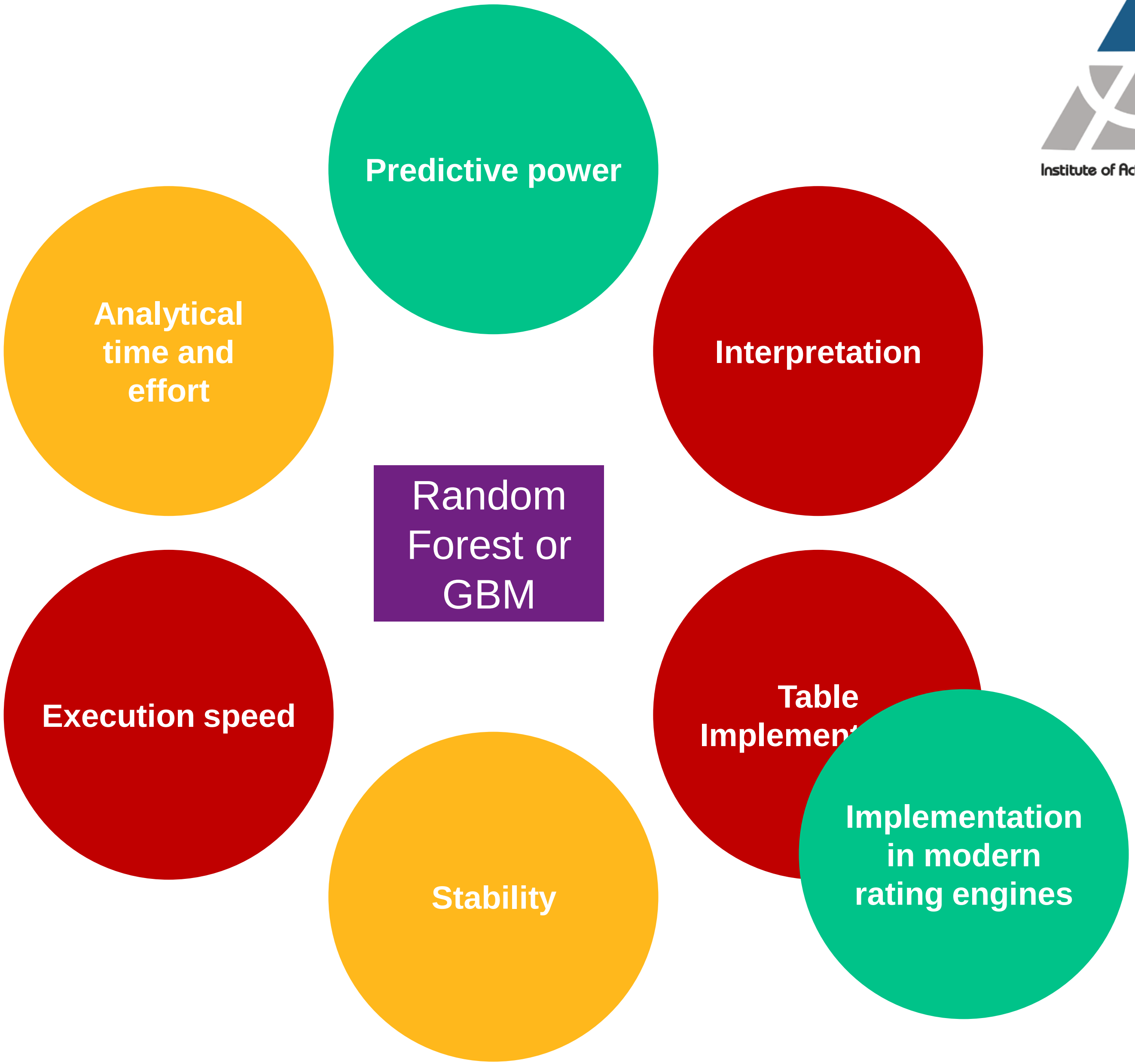
Support Vector Machines

Ridge Regression

Focus on Random Forests & Gradient Boosting Machines

Gradient Boosting Machines

Random Forests


Predictive power
Analytical time and effort
Interpretation
Random Forest or GBM
Execution speed
Table
Implemen
Implementation in modern rating engines
Stability

[redacted]

Regression-based methods

What are these machine learning methods?

Ensembles	Classification Trees	"Earth"	Regression Trees	Gradient Boosting Machines
K-nearest Neighbors	Elastic Net	Neural Networks	Naïve Bayes	Random Forests
K-Means Clustering	Principal Components Analysis	Lasso	Support Vector Machines	Ridge Regression

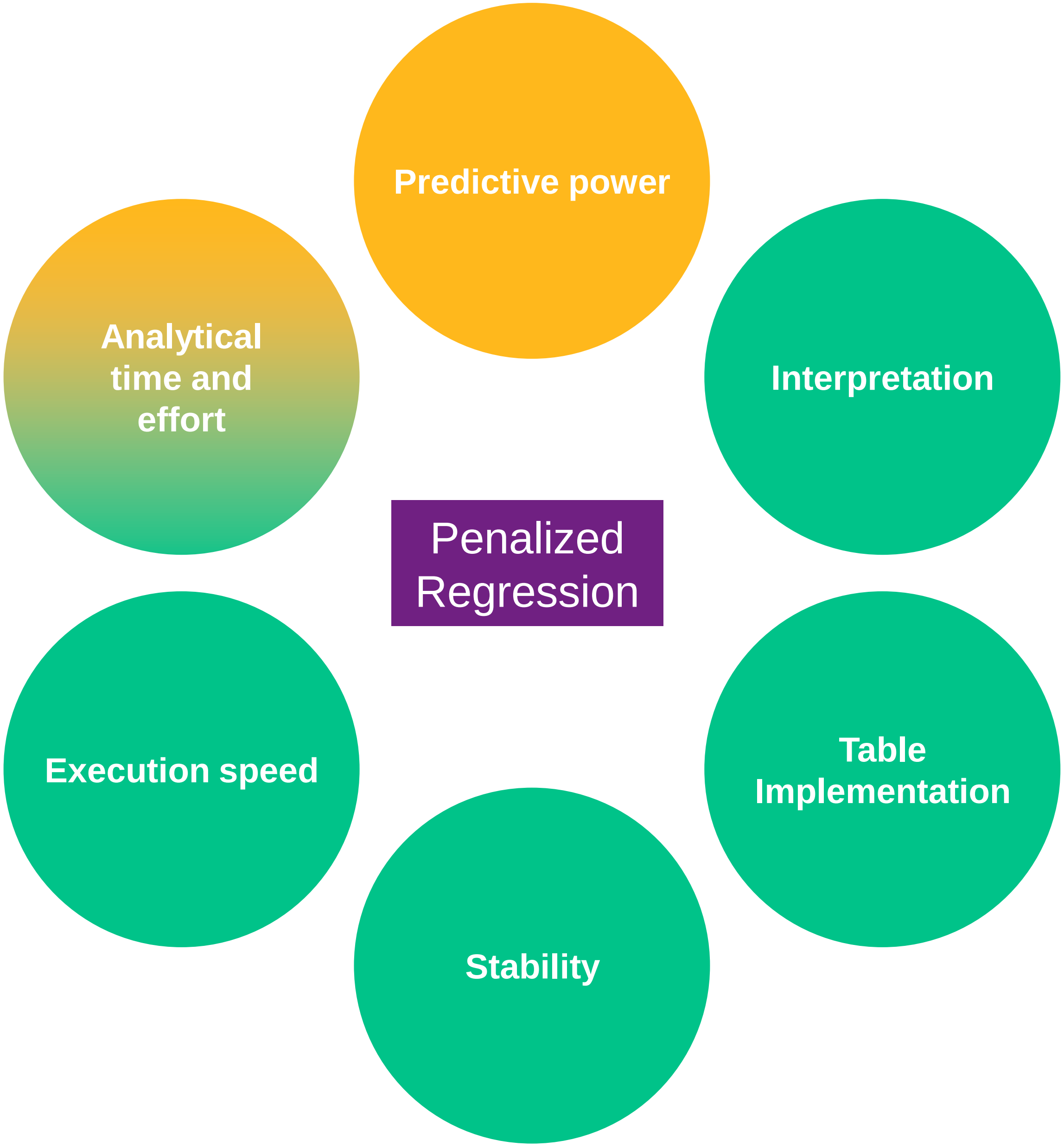
Focus on Penalized Regression

Elastic Net

Lasso

Ridge Regression

Institute of Actuaries of India



[redacted]

Pricing applications

Institute of Actuaries of India

Practical applications of machine learning methods in pricing

Direct ↔ **Indirect**

Unconstrained Models

Incorporating Nonlinearity

Interaction Detection

Peril Grouping

Feature Selection

Rating Factor Selection

Pricing Accuracy

Enhanced MI

[redacted]

Institute of Actuaries of India

How are insurers using machine learning?

Insights from a North American survey

Institute of Actuaries of India

How the North American market is doing with machine learning

Top applications insurers plan to use two years from now for AI and machine learning

	Actual for 2017	Expected for 2020 (in 2017)	Actual for 2020	Expected for 2021
Build risk models for better decision making	13%	44%		
Reduce time spent by humans	11%	49%		
Better understand risk drivers	21%	44%		
Identify cases that pose higher risk	11%	46%		
Augment human-performed underwriting	7%	37%		
Identify patterns of fraudulent claims	9%	39%		
Identify bottlenecks in claim processes/Process claims more efficiently	3%	30%		

How the North American market is doing with machine learning

Top applications insurers plan to use two years from now for AI and machine learning

	Actual for 2017	Expected for 2020 (in 2017)	Actual for 2020	Expected for 2021
Build risk models for better decision making	13%	44%	26%	
Reduce time spent by humans	11%	49%	22%	
Better understand risk drivers	21%	44%	20%	
Identify cases that pose higher risk	11%	46%	14%	
Augment human-performed underwriting	7%	37%	7%	
Identify patterns of fraudulent claims	9%	39%	17%	
Identify bottlenecks in claim processes/Process claims more efficiently	3%	30%	7%	

How the North American market is doing with machine learning

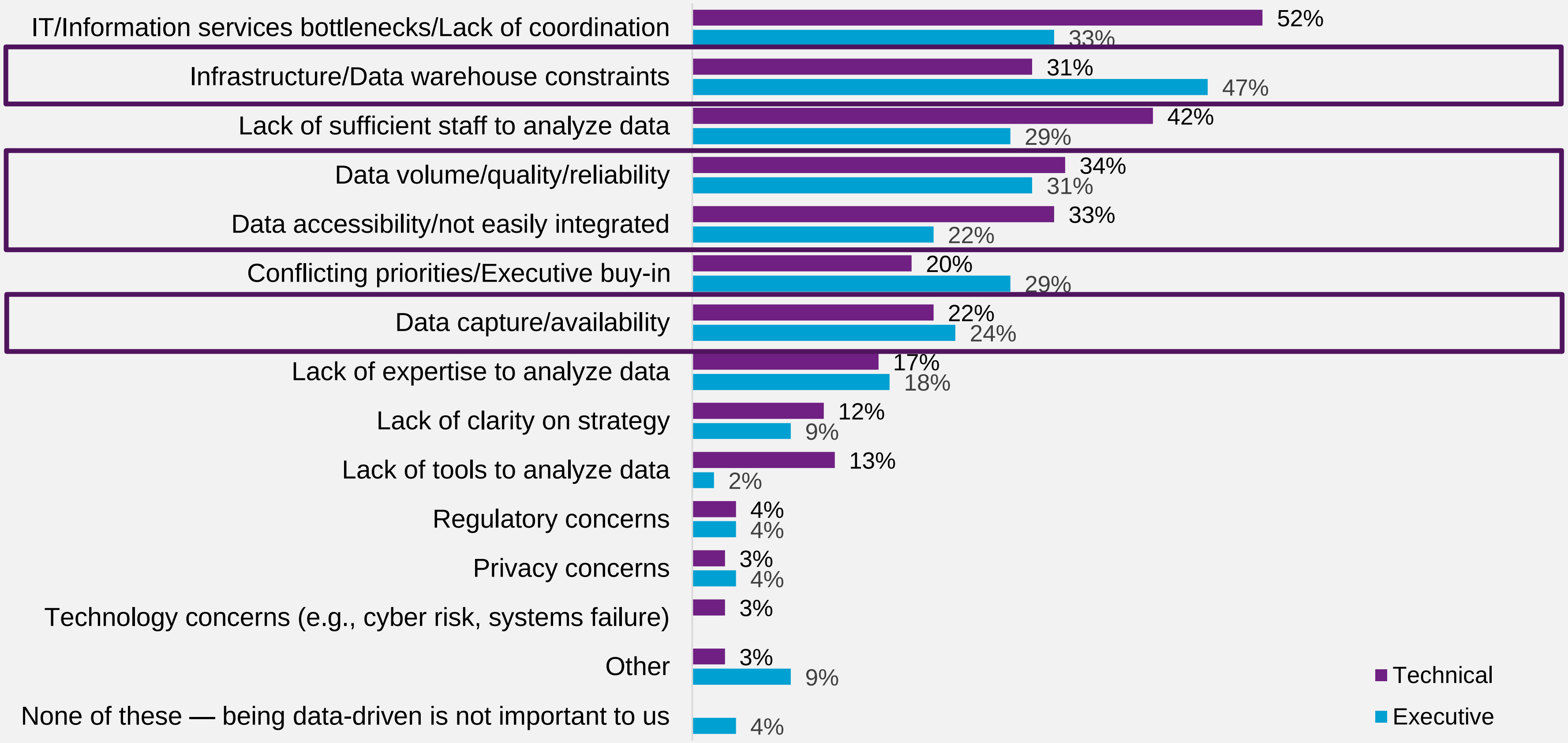
Institute of Actuaries of India

Top applications insurers plan to use two years from now for AI and machine learning

	Actual for 2017	Expected for 2020 (in 2017)	Actual for 2020	Expected for 2021
Build risk models for better decision making	13%	44%	26%	60%
Reduce time spent by humans	11%	49%	22%	60%
Better understand risk drivers	21%	44%	20%	56%
Identify cases that pose higher risk	11%	46%	14%	50%
Augment human-performed underwriting	7%	37%	7%	47%
Identify patterns of fraudulent claims	9%	39%	17%	47%
Identify bottlenecks in claim processes/Process claims more efficiently	3%	30%	7%	43%

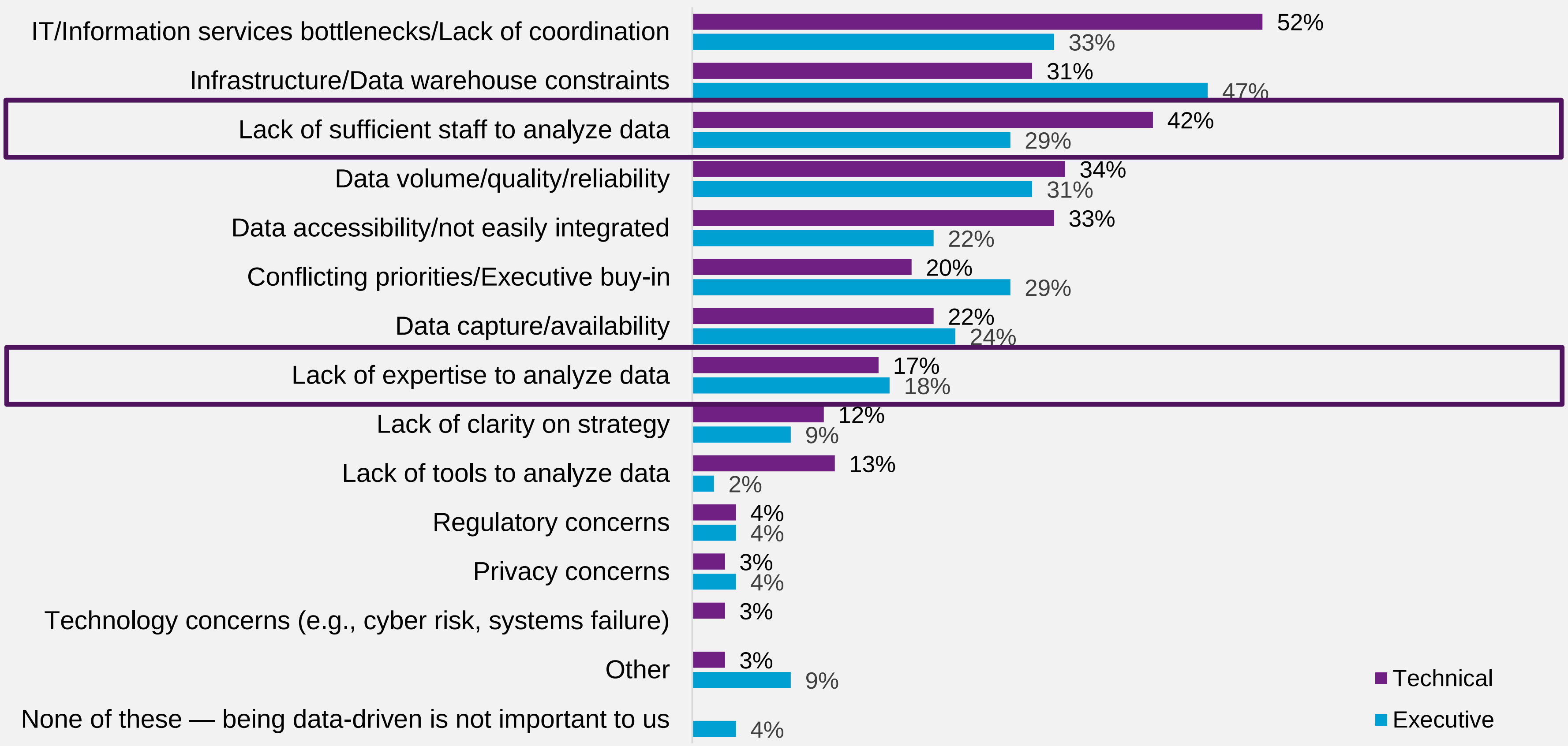
Institute of Actuaries of India

What are the three biggest challenges preventing your company from becoming more data driven?



Base: Total respondents (technical n = 77, executive n = 45)

What are the three biggest challenges preventing your company from becoming more data driven?

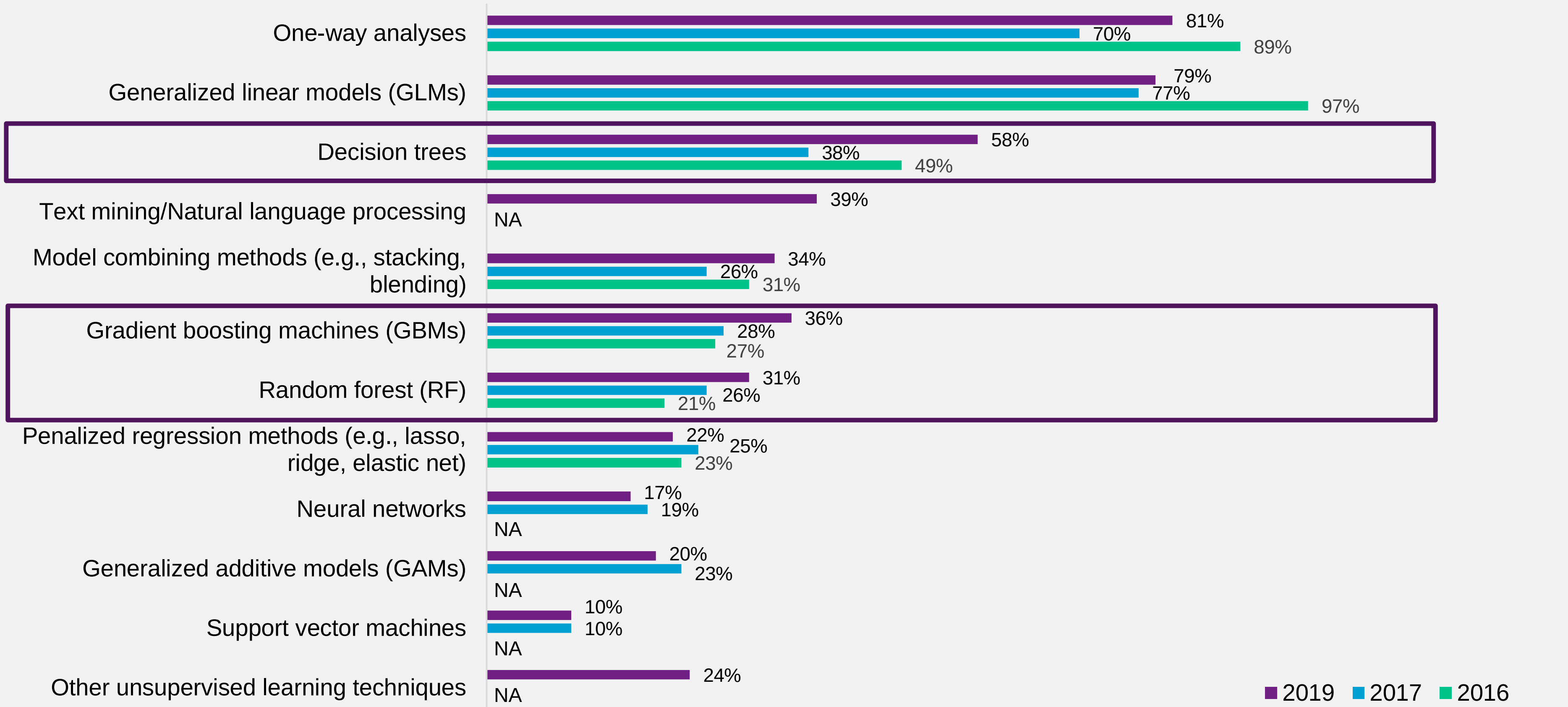


Base: Total respondents (technical n = 77, executive n = 45)

Institute of Actuaries of India

How the North American market is doing with machine learning

Methods used

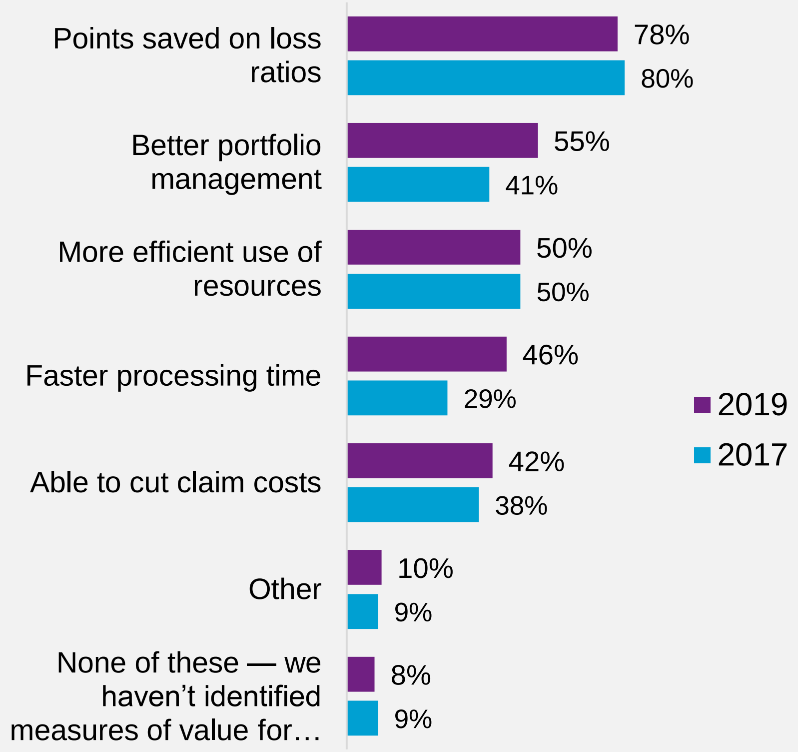


Base: Technical survey respondents using advanced analytics (n = 76 in 2020, n = 69 in 2017, n = 62 in 2016)

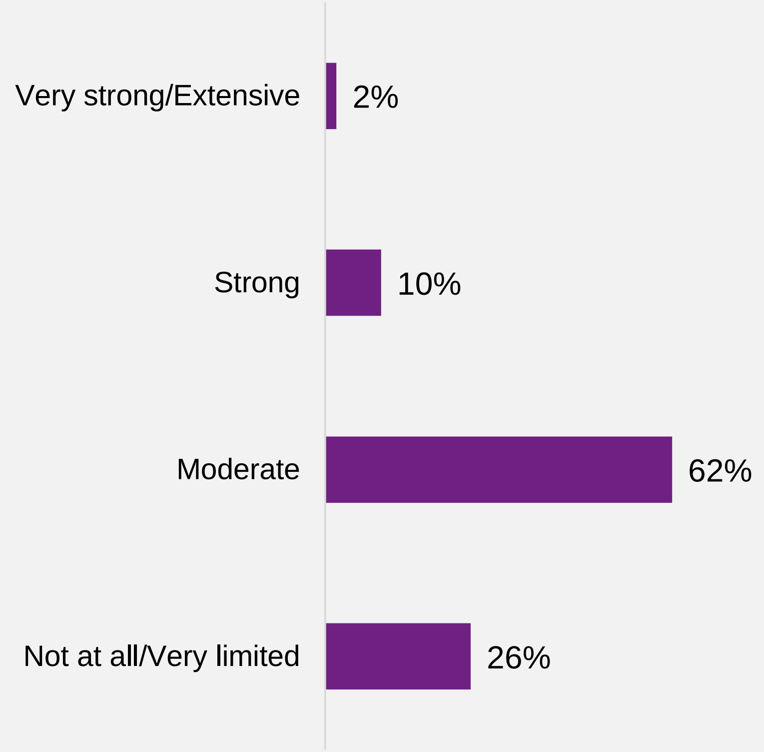
How do you determine the value of your predictive models?
How well understood are your predictive models by those who need to use them, outside of the modeling team?

Institute of Actuaries of India

Measures used to determine value of predictive models

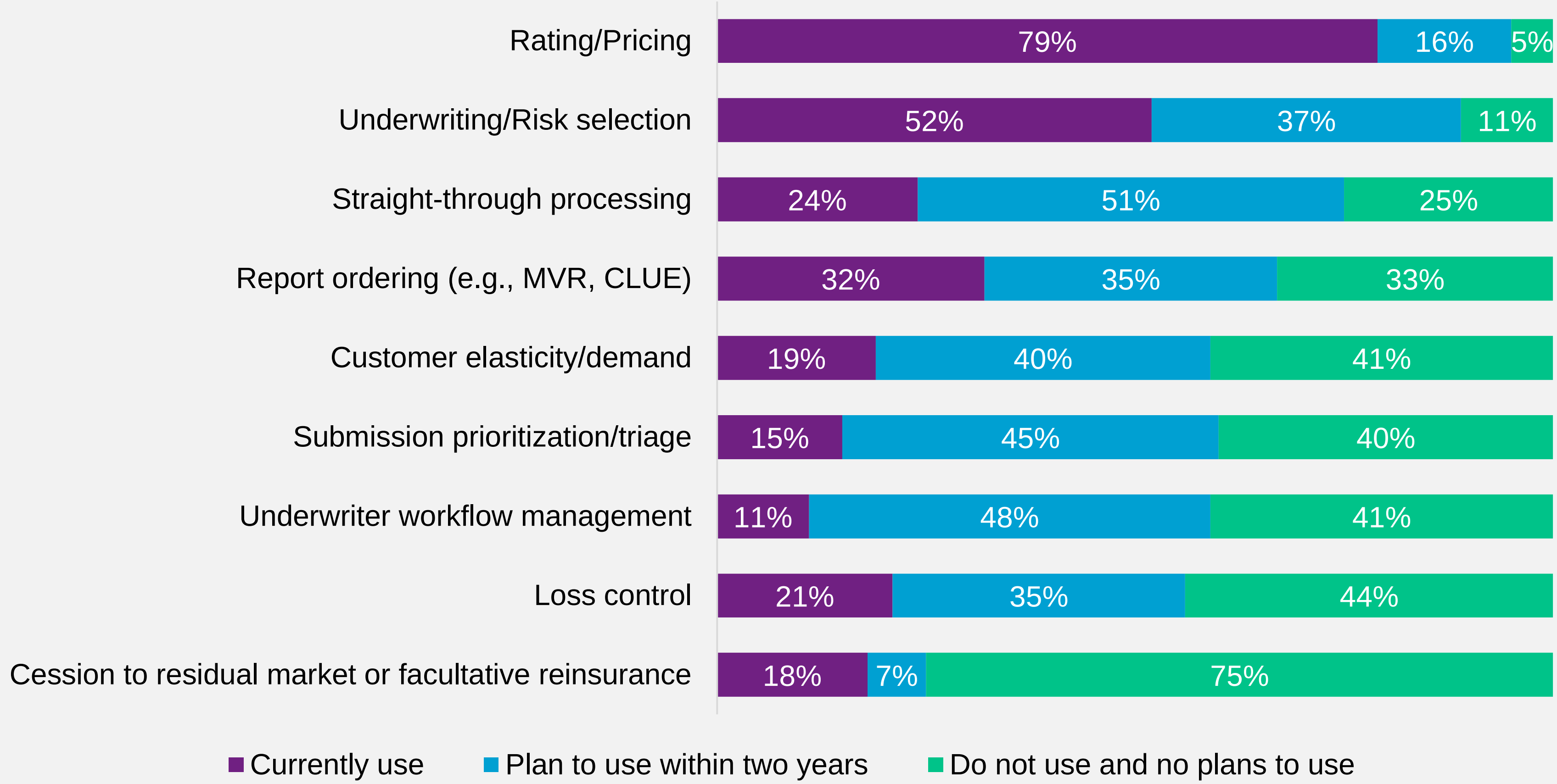


Level of understanding of predictive models outside of the modeling team



Base: Total respondents using advanced analytics (n = 113 in 2020, n = 56 in 2017)

For which aspects of underwriting/pricing does your company group currently use or plan to use advanced analytics?



Base: Total respondents (n = 122)

[redacted]

Adoption and use of machine learning in APAC and India

Classification of Indian Market

BEHIND THE MARKET

⓾Few companies are still struggling with legacy issues and have old methods of carrying out core insurance work

- **AT MARKET**

•Companies making limited use of AI/ML techniques for:

•1. Actuarial

•2. Claims Management/ Fraud Analytics

•3. Underwriting

•4. Sales

ADVANCED

Companies using AI/ML techniques for:

1. Actuarial
2. Claims Management/ Fraud Analytics
3. Underwriting
4. Sales
5. Marketing
6. Customer Support
7. Human Resources

GLM and ML in APAC

Adopting of GLM and Machine Learning in APAC

Education and skillsets

Actuary vs data scientist

GLM or ML?

Predictive power

Interpretation

Table Implementation

Implementation in modern rating engines

Stability

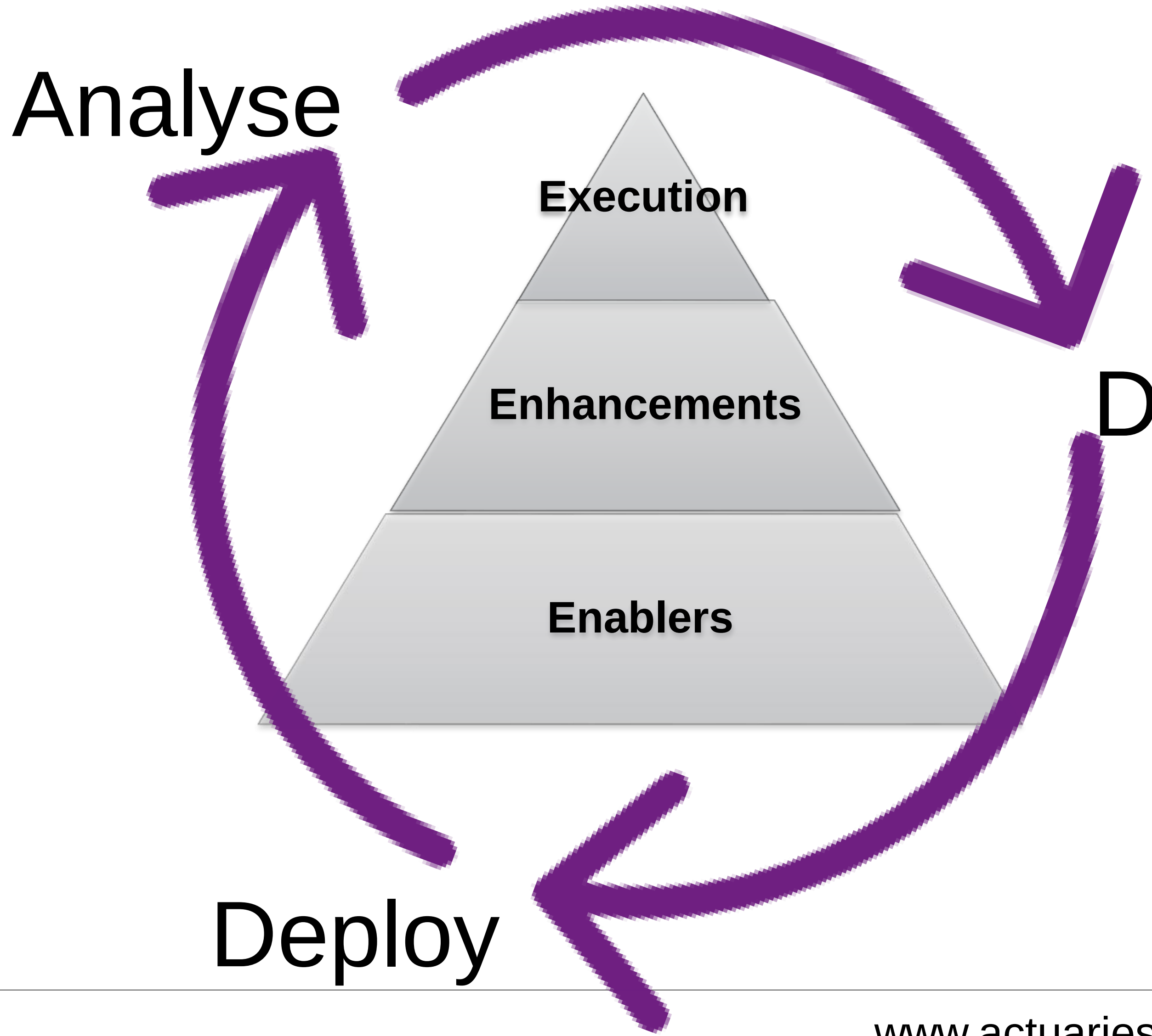
Execution speed

Analytical time and effort

Decision tree
Random Forest
Elastic net
GBM
GLM

Start your journey today

Institute of Actuaries of India



[redacted]

Summary and Q&A

A more complete summary

Institute of Actuaries of India

Institute of Actuaries of India

Machine Learning in Pricing

Conclusions

- There are many forms of ML models
- New data and feature/response engineering generally add more value than new methods BUT we need to continuously explore which methods work on which problems
- Traditional measures of prediction value may not reflect applications in insurance
- And it's not all about predictive power anyway – other criteria are important

- GBMs and Random Forests can provide predictive lift benefits by capturing higher order effects … BUT
  - Can you cope with not seeing the model and instead use broad diagnostics?
  - Effort is required to expose/understand higher order effects in an expeditious manner
  - How will business leaders and regulators respond to these methods?
  - Can you file and deploy results based on these models?

Institute of Actuaries of India

Machine Learning in Pricing

Conclusions

- Penalized regression can aid in factor selection decisions and may in fact be a good method in its own right – particularly when the modeler has less of a "feel" for the data
- Machine learning in pricing is not all about improving predictive power. Consider:
  - Fast investigation of new data
  - Quick assessment and response of emerging experience

Machine Learning beyond Pricing

Conclusions

- Machine Learning is becoming established within insurance analytics
- It opens up a broader set of problems to analytics, and offers a broader tool set for familiar problems
- There's opportunity to reveal actionable, first-order insights in applications to which analytics have not been deployed previously
- We expect use of Machine Learning to continue to grow- in Pricing and beyond

WillisTowersWatson