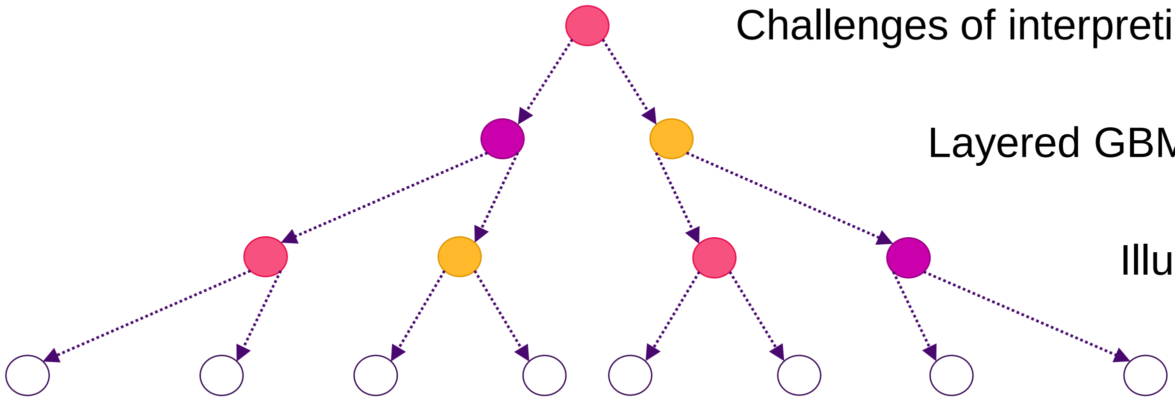
Introduction to Layered GBM

Made for insurance, interpretable by design

Liam McGrath and Justin Milam

November 9, 2022

Agenda



Challenges of interpreting traditional GBMs

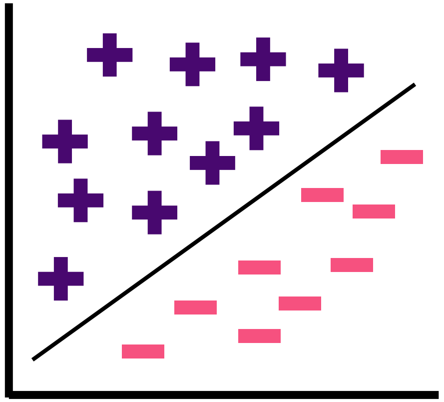
Layered GBMs: inherently interpretable

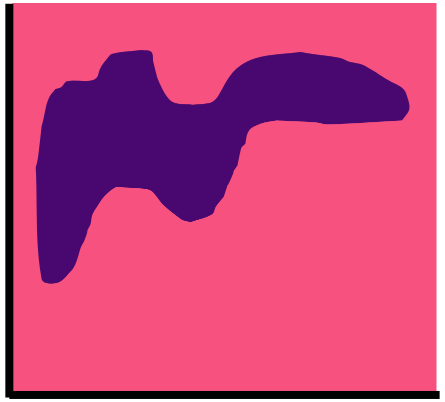
Illustrative example of LGBM

Case Studies

Challenges with Traditional GBMs

Interpreting models





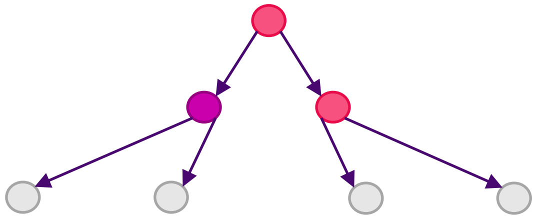
Question 1: Could you model this relationship with a GLM?

Question 2: Could you explain this to a regulator?

Traditional GBMs

Start with a decision tree

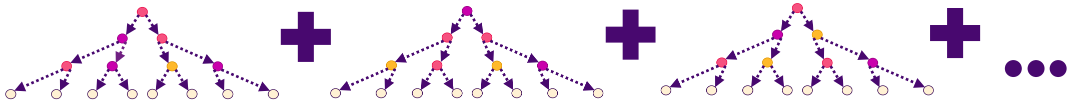
1. Consider every possible way of splitting our data, choose “best” split (greedy algorithm)
2. Assign the average response value of observations in a node as the node’s parameter value
3. For each subsequent node, repeat from 1



Not very predictive

Highly unstable

Boost many decision trees



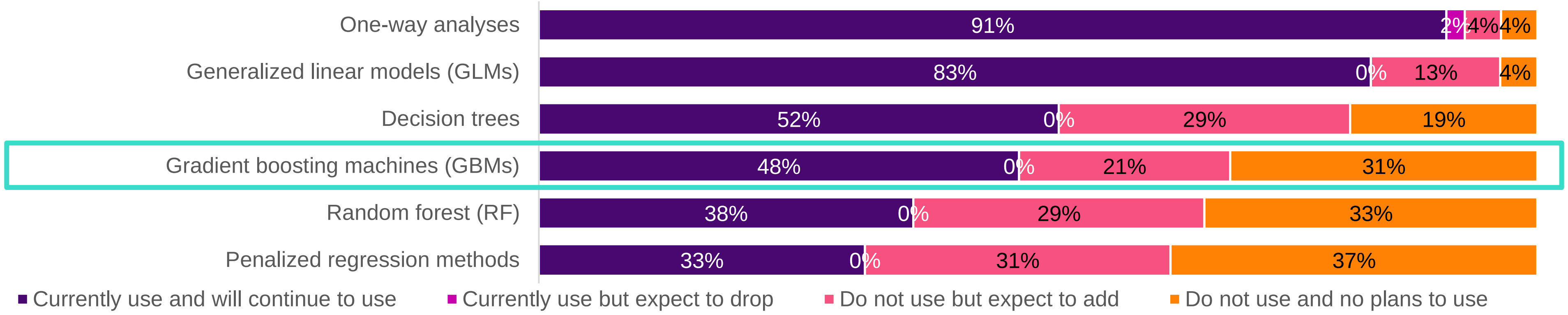
Very predictive

Stable

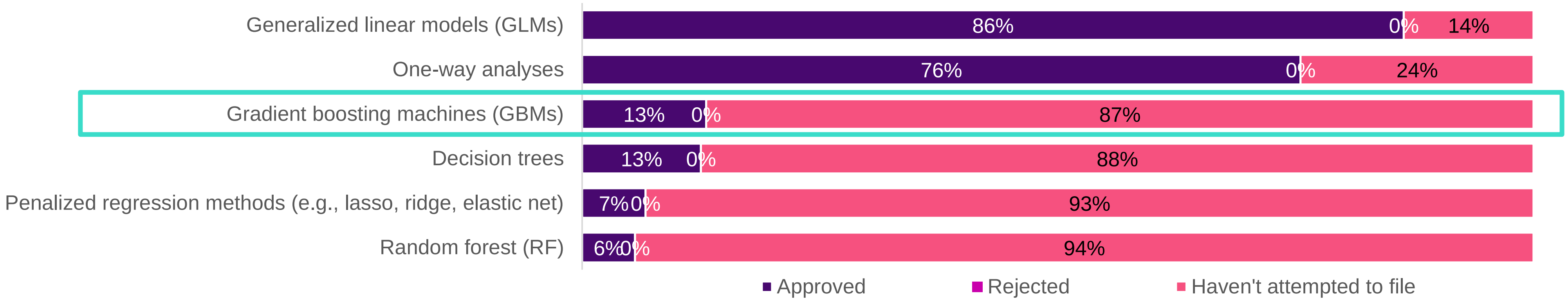
Captures non-linear effects

Reduced interpretability

Which techniques do you currently use and which do you plan to use in the next two years?

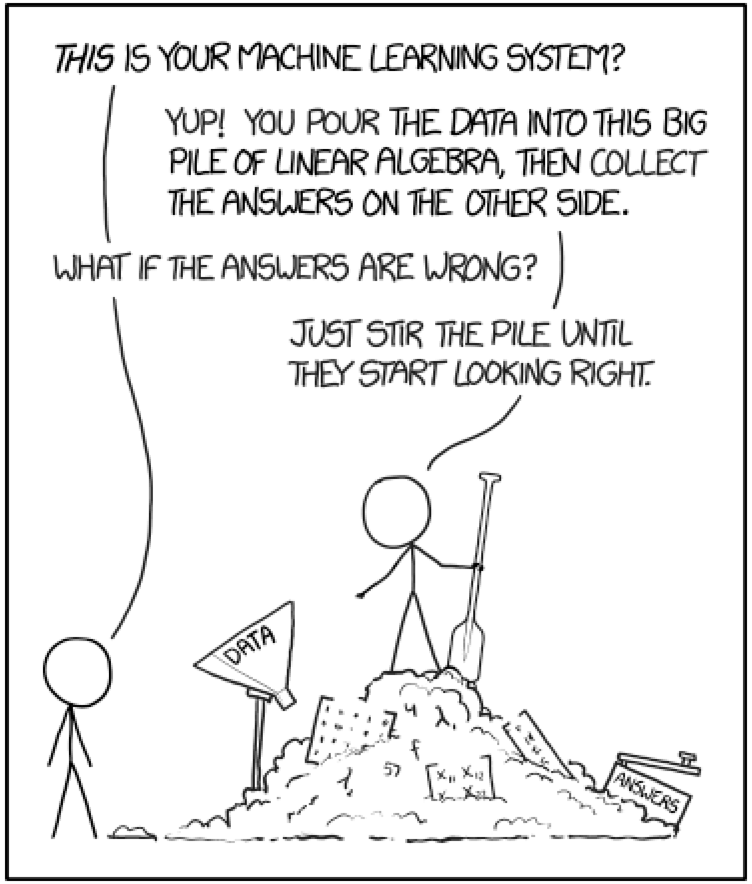


Which of these techniques have been approved by regulators in rate filings?



Source: *2021 P&C Advanced Analytics Survey*, Willis Towers Watson

Why should we care about interpretability?



Source: https://xkcd.com/1838/

External

- Regulatory compliance
- Ethical standards
- Policyholder retention

Internal

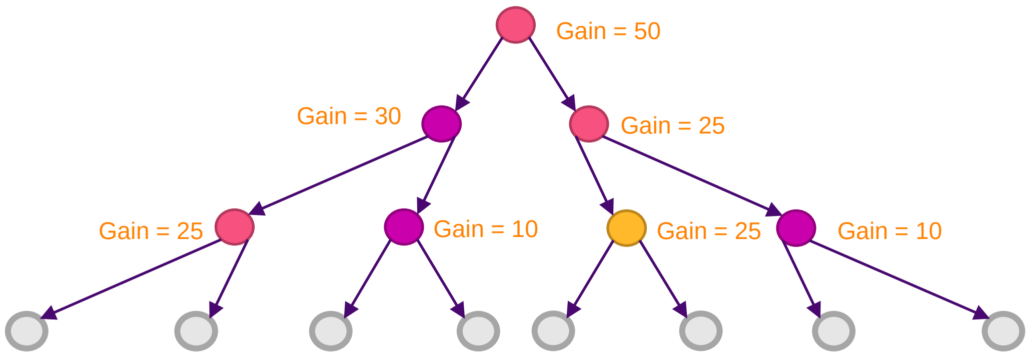
- Domain knowledge
- Robust models
- Management approval
- Informed decisions
- Debugging

Traditional GBMs

Interpretation: Variable Importance

It is unknown whether the gain of splits beyond the first level are due to main effects or interactions with higher splits.

1. For all splits involving a feature, calculate the gain (loss reduction) from the split
2. Add gains across all splits and trees



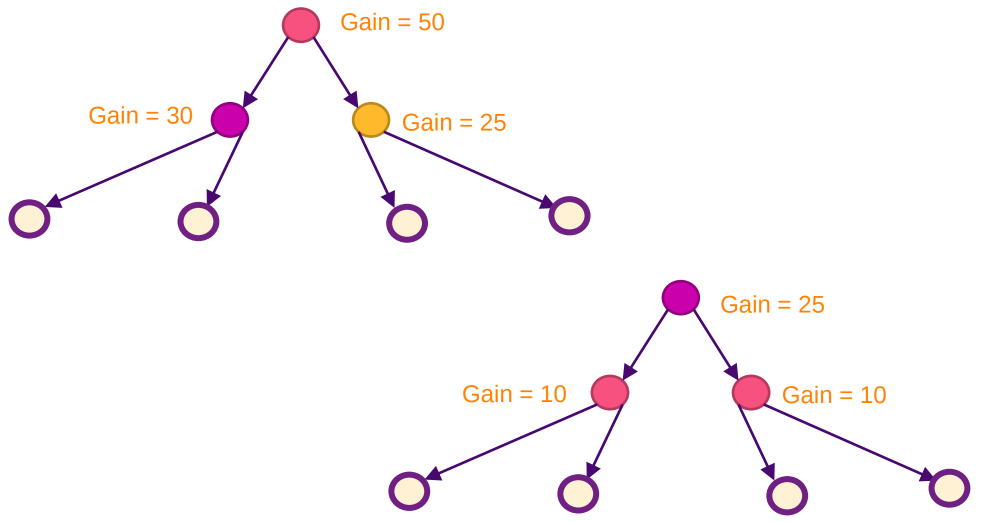
Feature	Total Gain	Gain from one-way effect	Interaction gain
	100	75 + ??	??
	50	??	??
	25	??	??

Improved* GBMs

Interpretation: Variable Importance

*In an ideal world

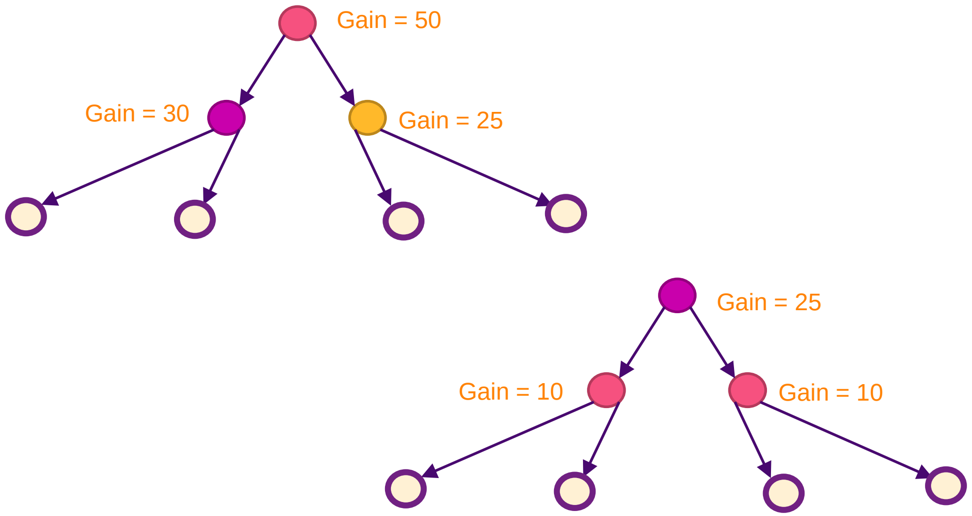
Imagine a GBM with one-way and interaction effects captured separately (layer 1 is main effects only, layer 2 is depth-2 interaction effects only, etc.).



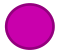
Feature	Total Gain	Gain from one-way effect	Gain from 2-way interaction
	70	50	20
	55	25	30
	25	0	25

Layered GBMs

Interpretation: Variable Importance

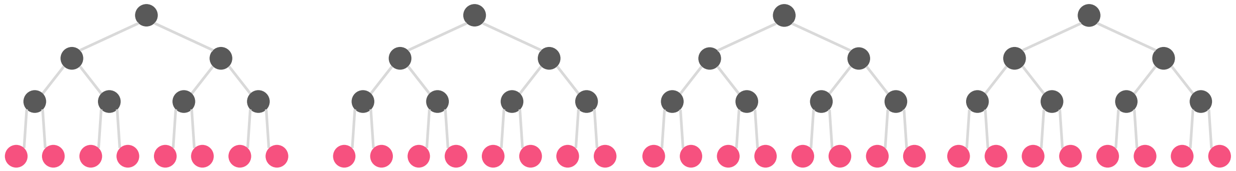


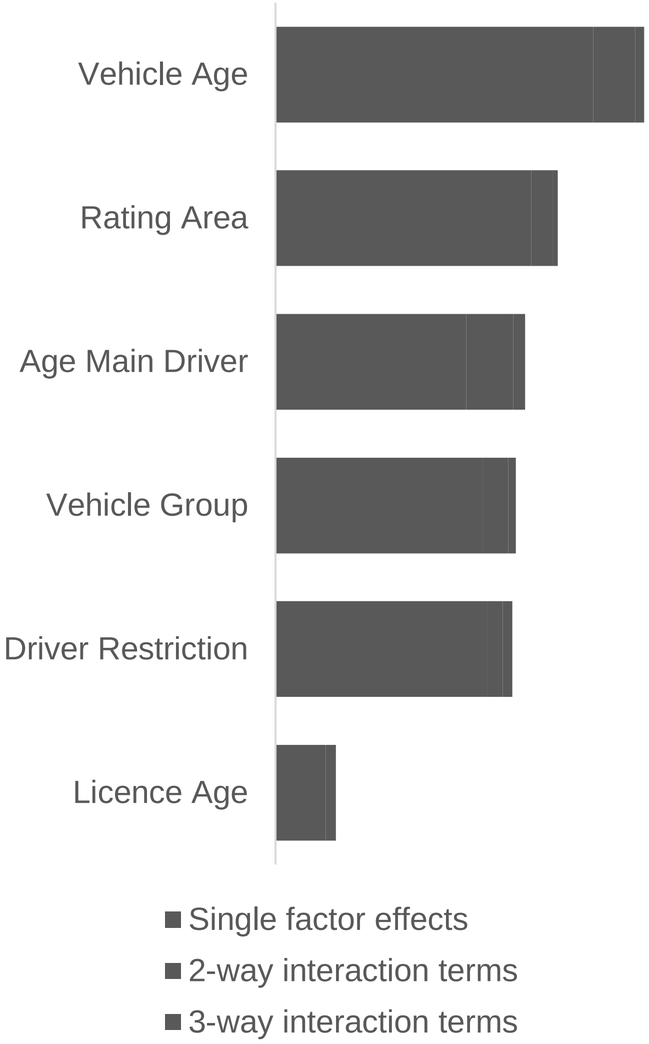
Imagine a GBM with one-way and interaction effects captured separately (layer 1 is main effects only, layer 2 is depth-2 interaction effects only, etc.).

Feature	Total Gain	Gain from one-way effect	Gain from 2-way interaction
	70	50	20
	55	25	30
	25	0	25

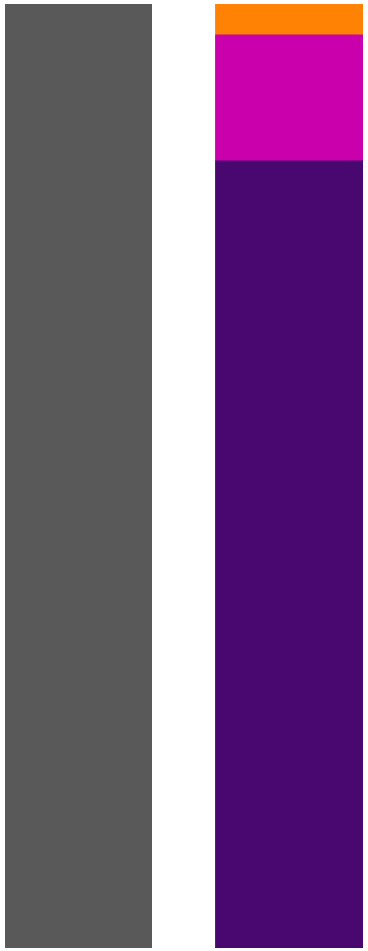
Introducing: Layered GBMs

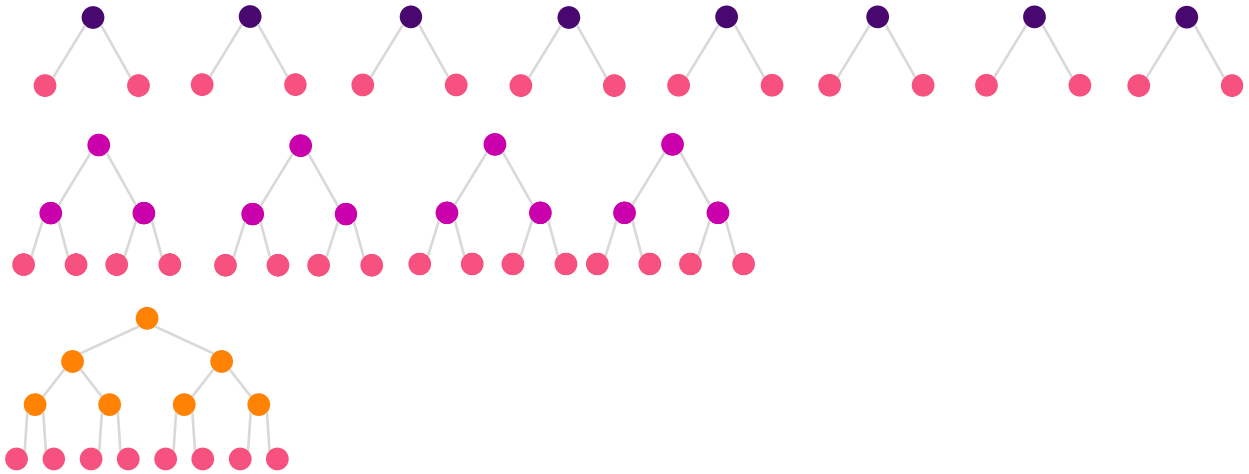
Traditional GBMs



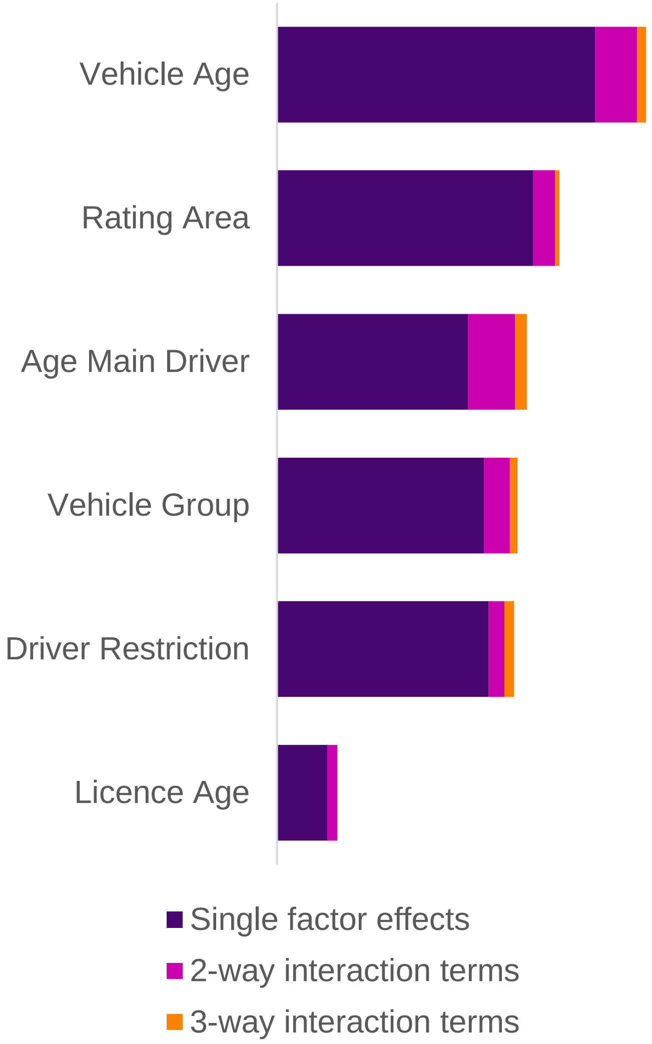


Layered GBMs





Overall predictiveness is the same, but with improved interpretability



LGBM Illustrative Example

Layered GBM Experiment

Experiment design:

- Two different “true” processes
  - No Interactions
  - Interactions

1. For each dataset, fit traditional GBM and Layered GBM
2. Compare factor importance
   - Which model tells us more?
3. Compare model fit
   - Is one model more predictive?

Process 1:

- Age of Main Driver
- BonusMalus
- Gender of Main Driver **Male ↑**
- Marital Status **Single ↑**
- Mileage
- Number of Drivers **1 ↑**
- Rating Area

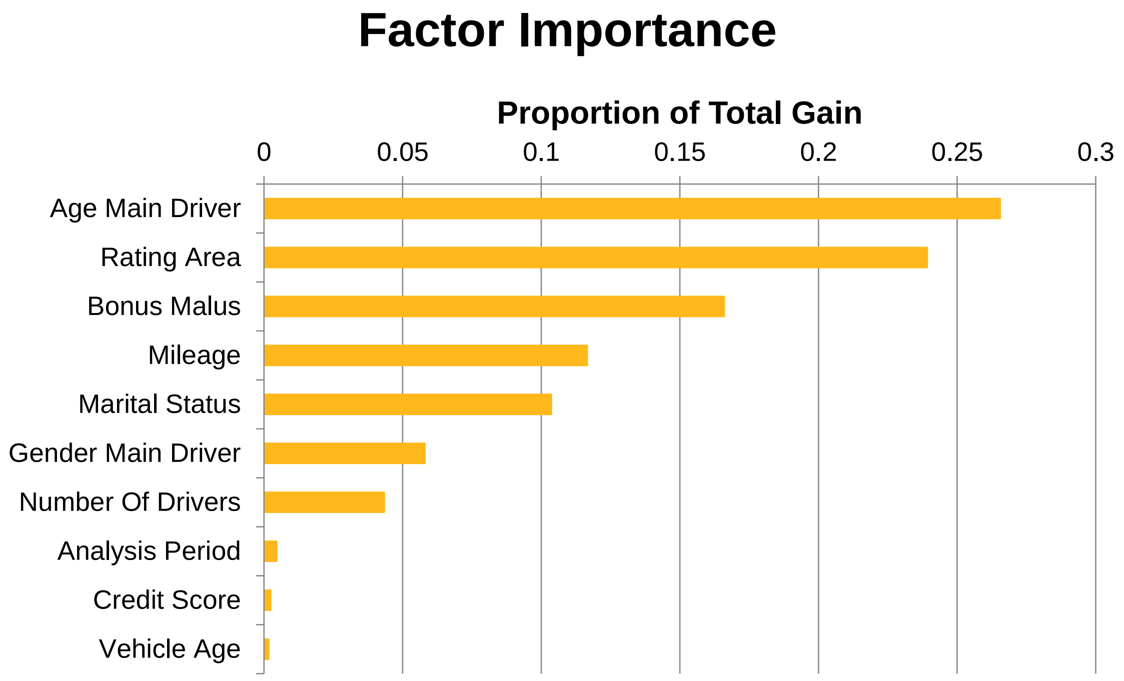
Process 2:

- Everything from Process 1
- Age x Credit **Young, Low Score ↑**
- Age x Marital x Gender **Young, Single, Male ↑**

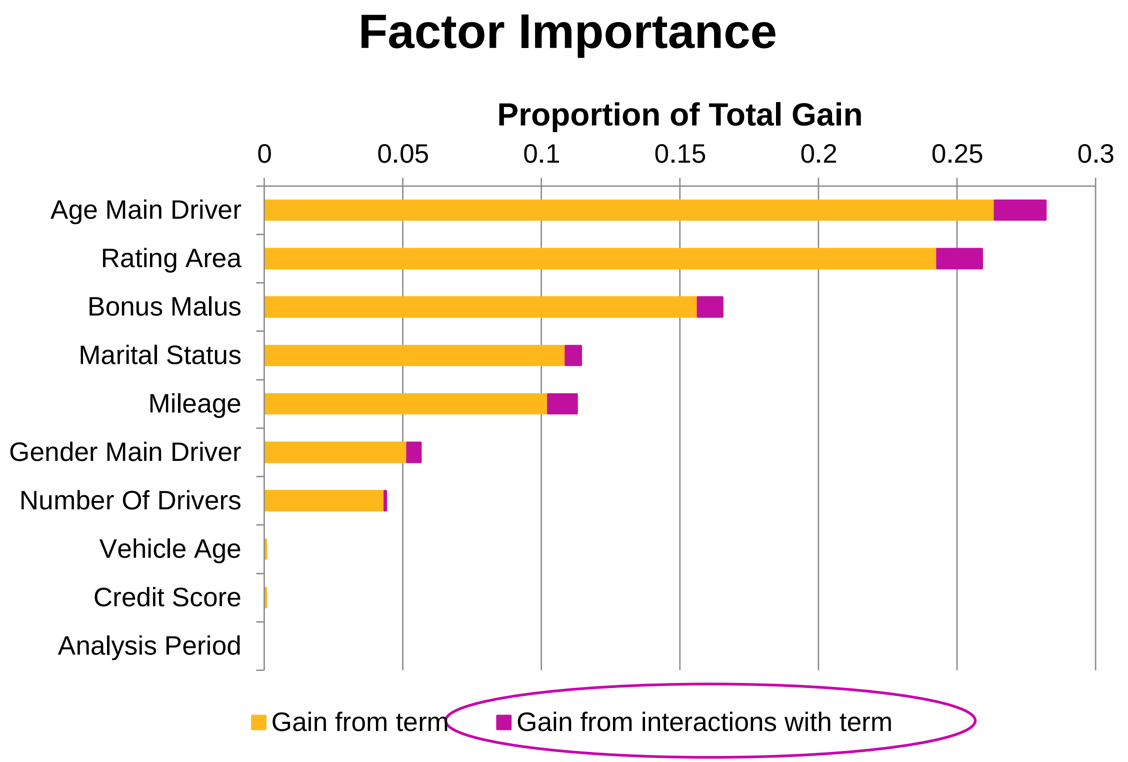
Experiment Results – Factor Importance

Process 1 (no interactions)

Traditional GBM

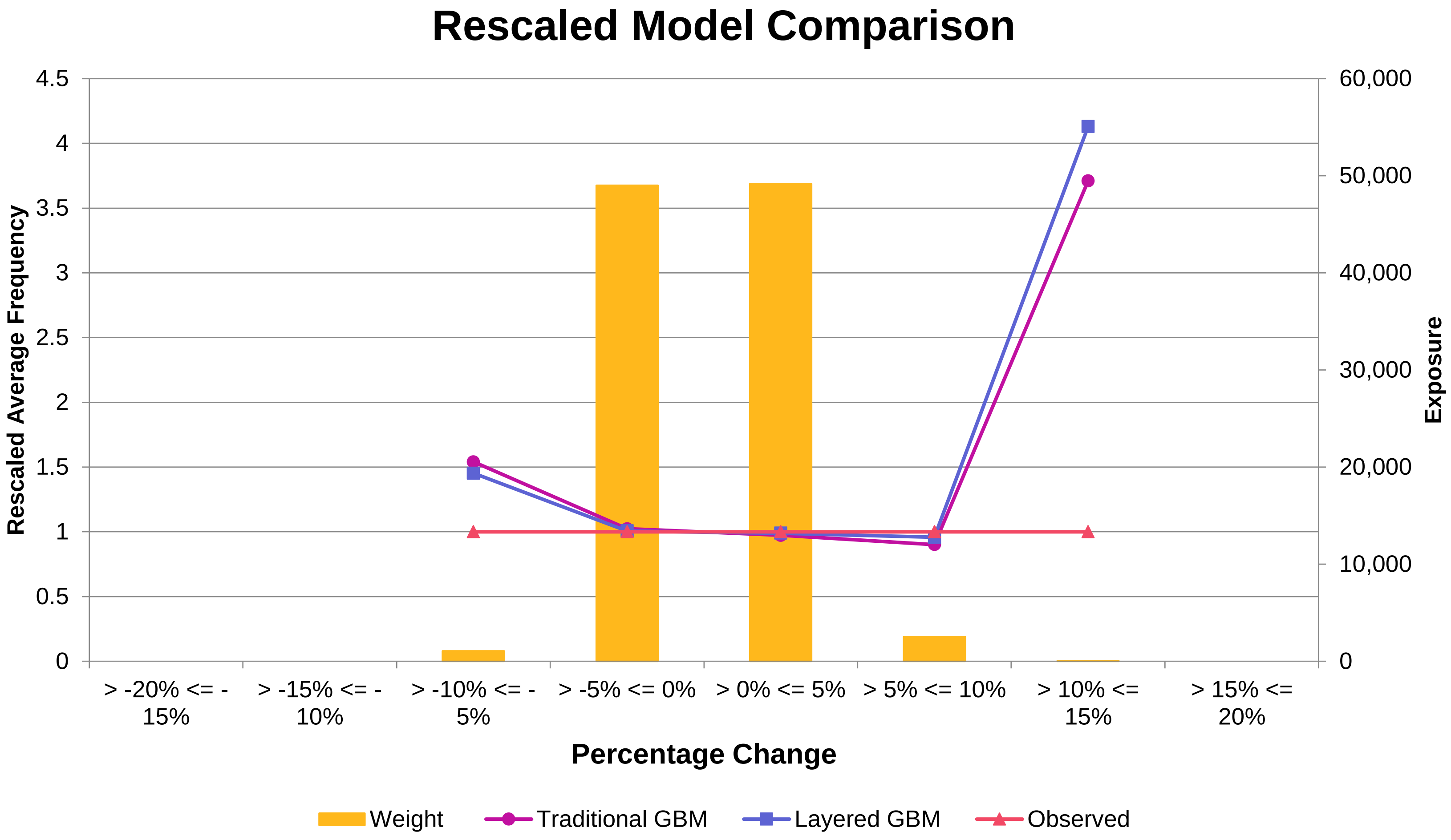


Layered GBM



Experiment Results – Model Comparison

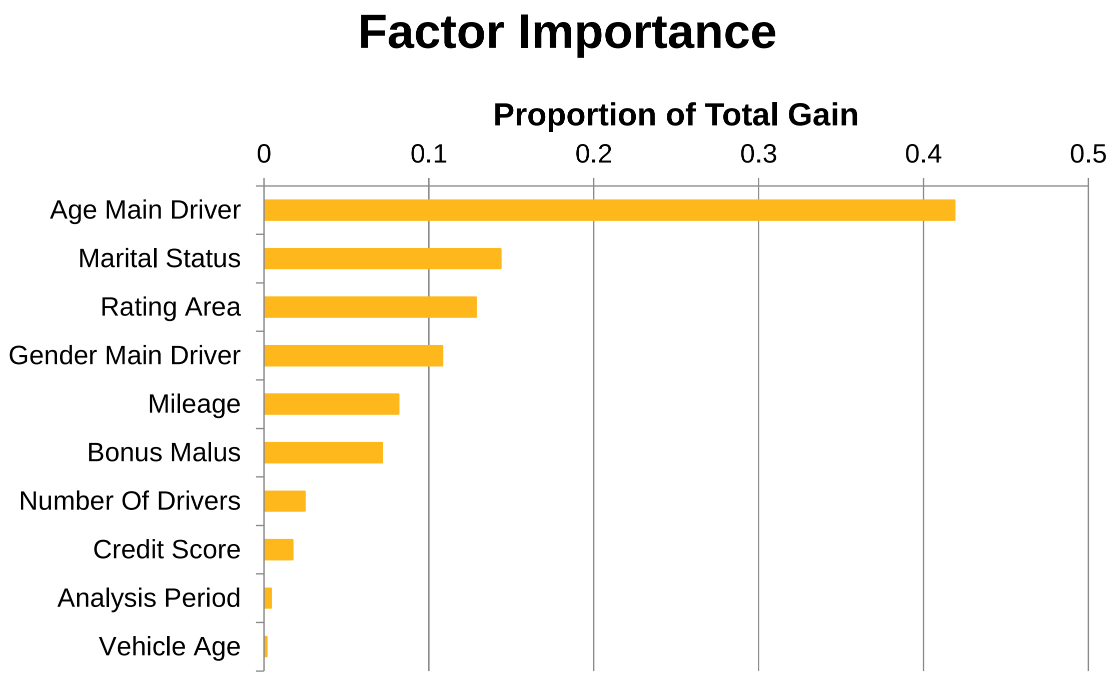
Process 1 (no interactions)



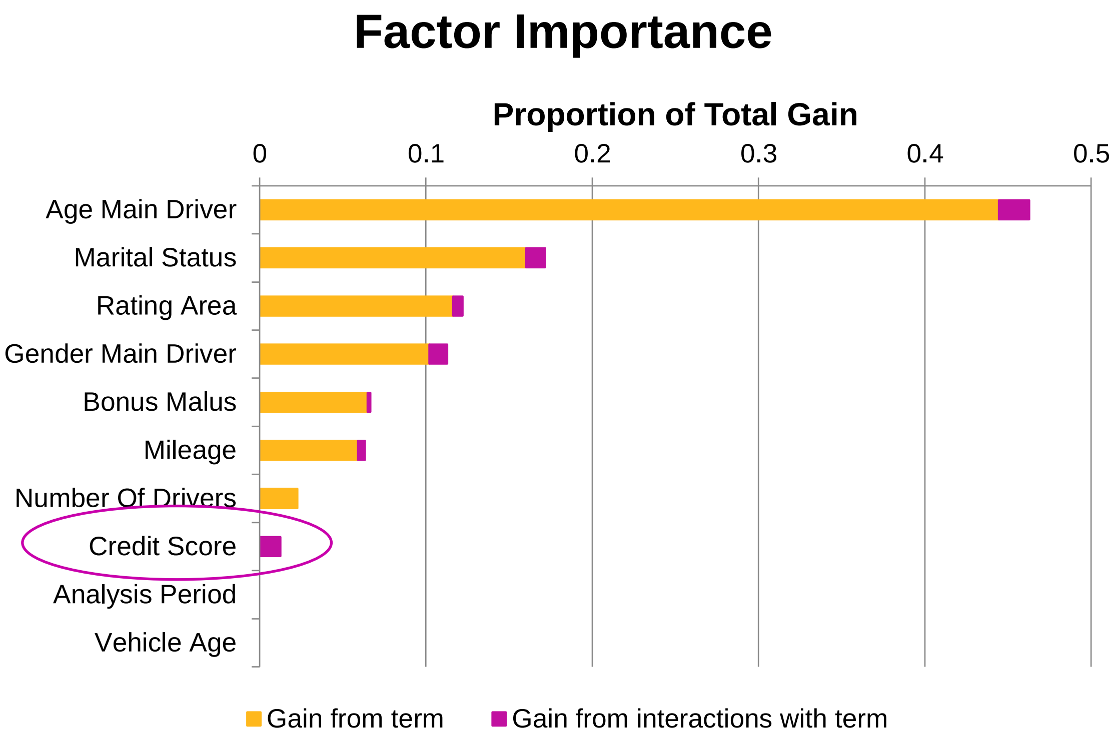
Experiment Results – Factor Importance

Process 2 (interactions)

Traditional GBM



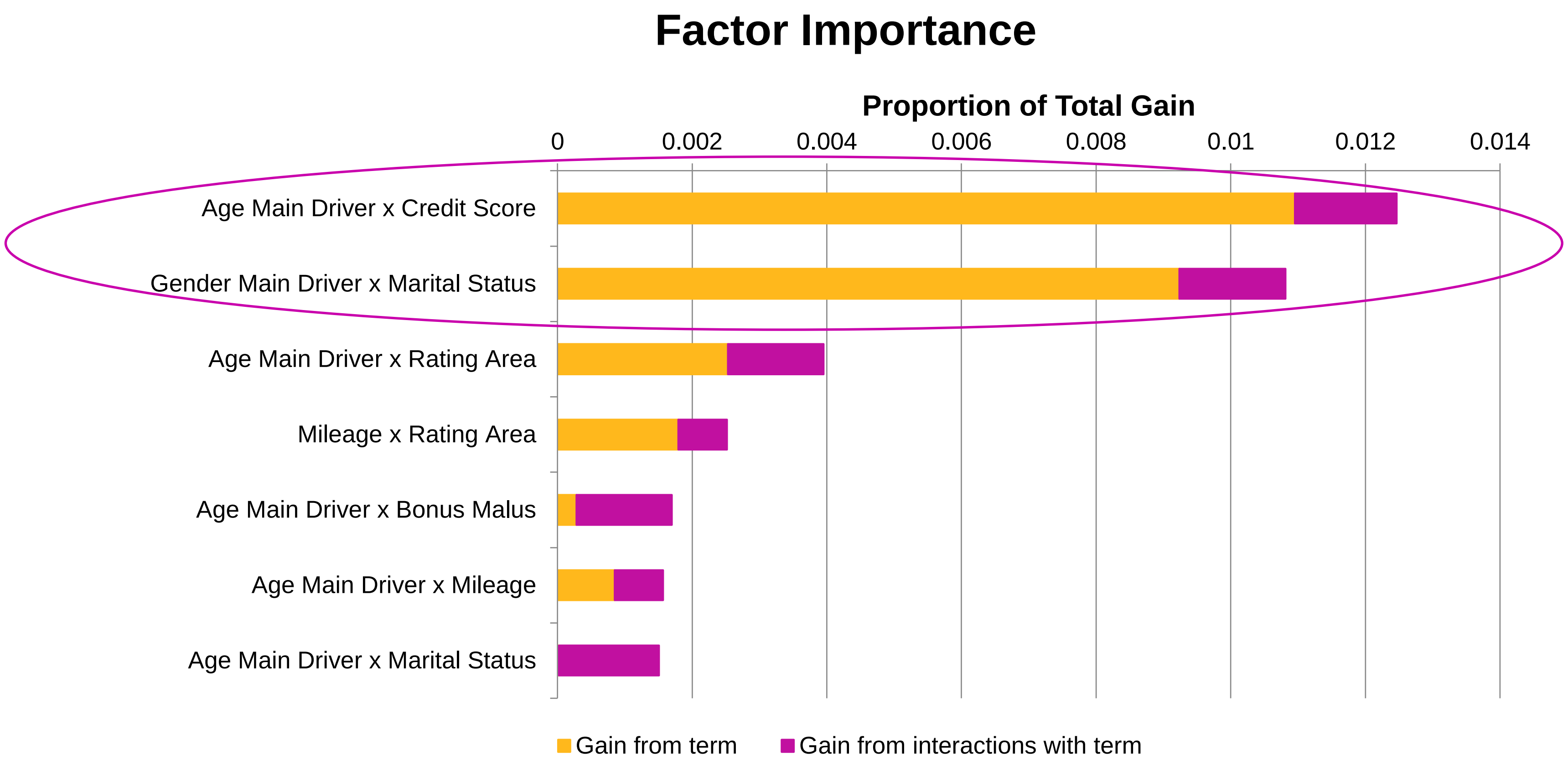
Layered GBM



Experiment Results – Factor Importance

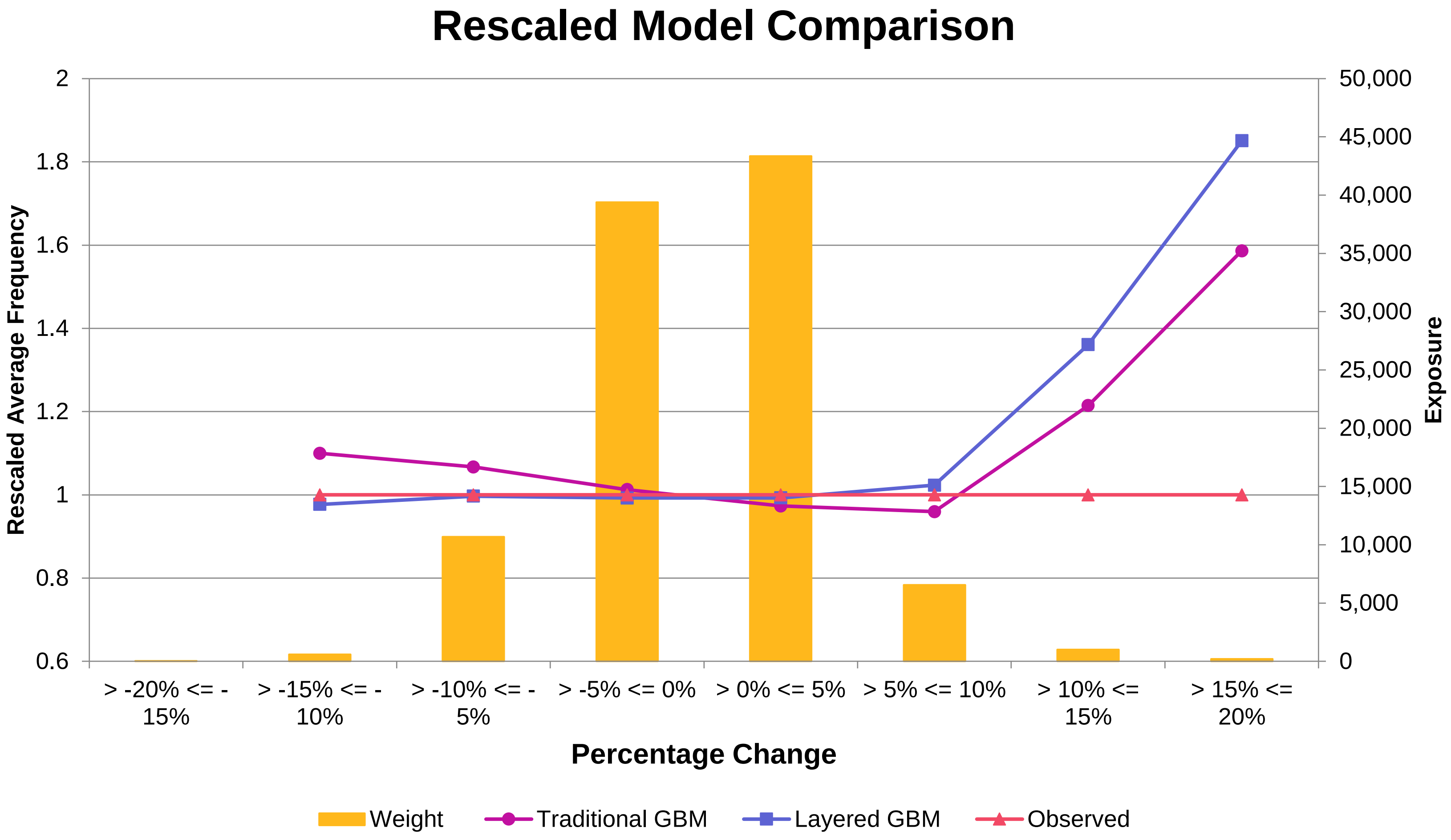
Process 2 (interactions)

Layered GBM 2-way Interactions



Experiment Results – Model Comparison

Process 2 (interactions)



LGBM: Looking Ahead

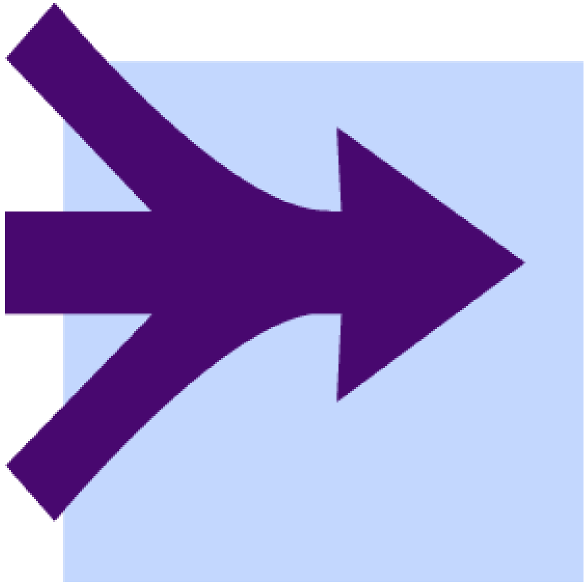
Interpretability techniques vs. Interpretable models

- Improved variable importance (covered here)
- Improved versions of PDPs
  - No need for a "representative" sample of data
  - Explicitly separate the marginal effects for each layer
- More useful guidance when building GLMs
  - Separately identify interactions

Case Studies

Use Case 1: Selecting factors to review

- You are doing a personal auto pricing review for the state of Minnesota.
- The product manager has asked you to review individual rating variables but you do not have time to review the entire rating algorithm.
- You decide to use a layered GBM to determine which variables should be reviewed.



Variables

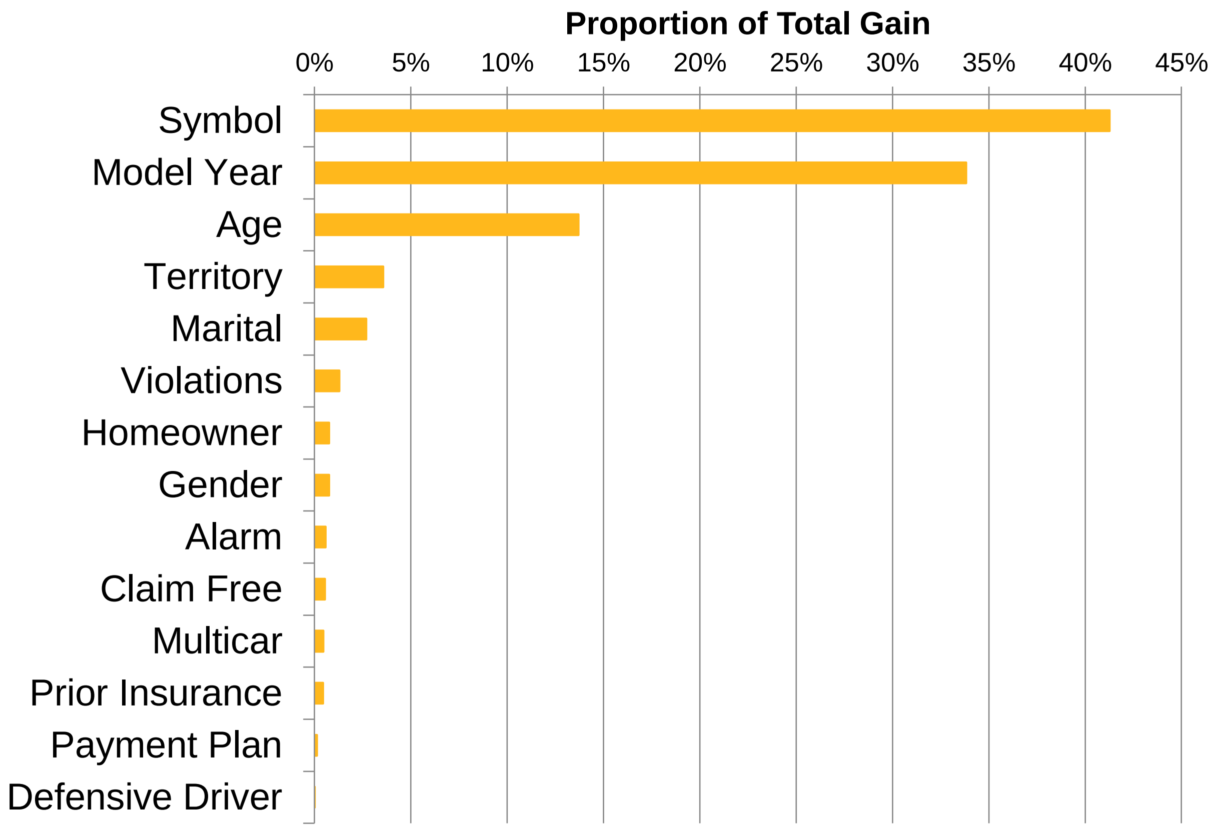
- Symbol
- Model Year
- Age
- Marital
- Violations
- Homeowner
- Alarm
- Claim Free
- Multicar
- Territory
- Gender
- Defensive Driver
- Prior Insurance Payment Plan

Use Case 1: Selecting factors to review

Proportion of Total Gain

Variable	Factor Importance
Symbol	41%
Model Year	34%
Age	14%
Territory	4%
Marital	3%
Violations	1%
Homeowner	1%
Gender	1%
Alarm	1%
Claim Free	1%
Multicar	0%
Prior Insurance	0%
Payment Plan	0%
Defensive Driver	0%

Factor Importance

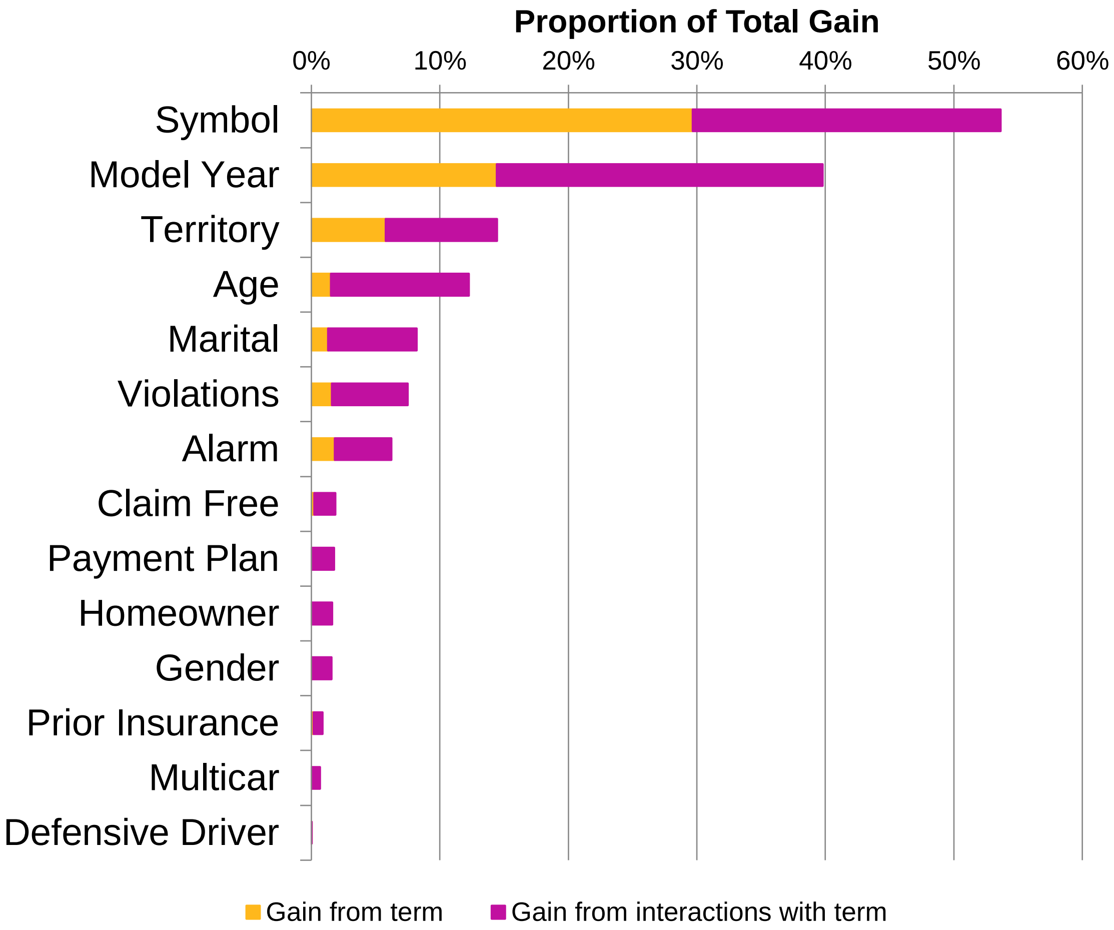


Use Case 1: Selecting factors to review

Proportion of Total Gain

Variable	Gain from term	Gain from interactions with term
Symbol	30%	24%
Model Year	14%	25%
Territory	6%	9%
Age	2%	11%
Marital	1%	7%
Violations	2%	6%
Alarm	2%	4%
Claim Free	0%	2%
Payment Plan	0%	2%
Homeowner	0%	2%
Gender	0%	2%
Prior Insurance	0%	1%
Multicar	0%	1%
Defensive Driver	0%	0%

Factor Importance

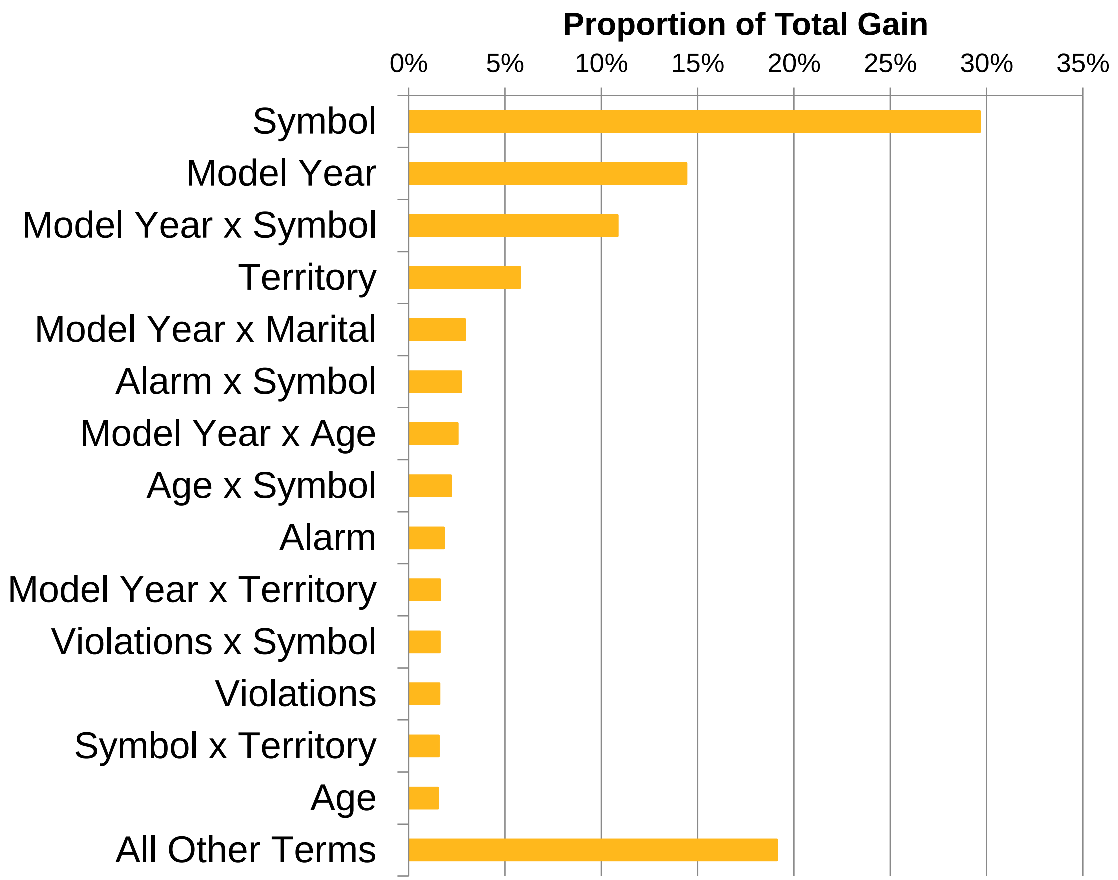


Use Case 1: Selecting factors to review

Proportion of Total Gain

Variable	Factor Importance
Symbol	30%
Model Year	14%
Model Year x Symbol	11%
Territory	6%
Model Year x Marital	3%
Alarm x Symbol	3%
Model Year x Age	3%
Age x Symbol	2%
Alarm	2%
Model Year x Territory	2%
Violations x Symbol	2%
Violations	2%
Symbol x Territory	2%
Age	2%
All Other Terms	19%

Factor Importance

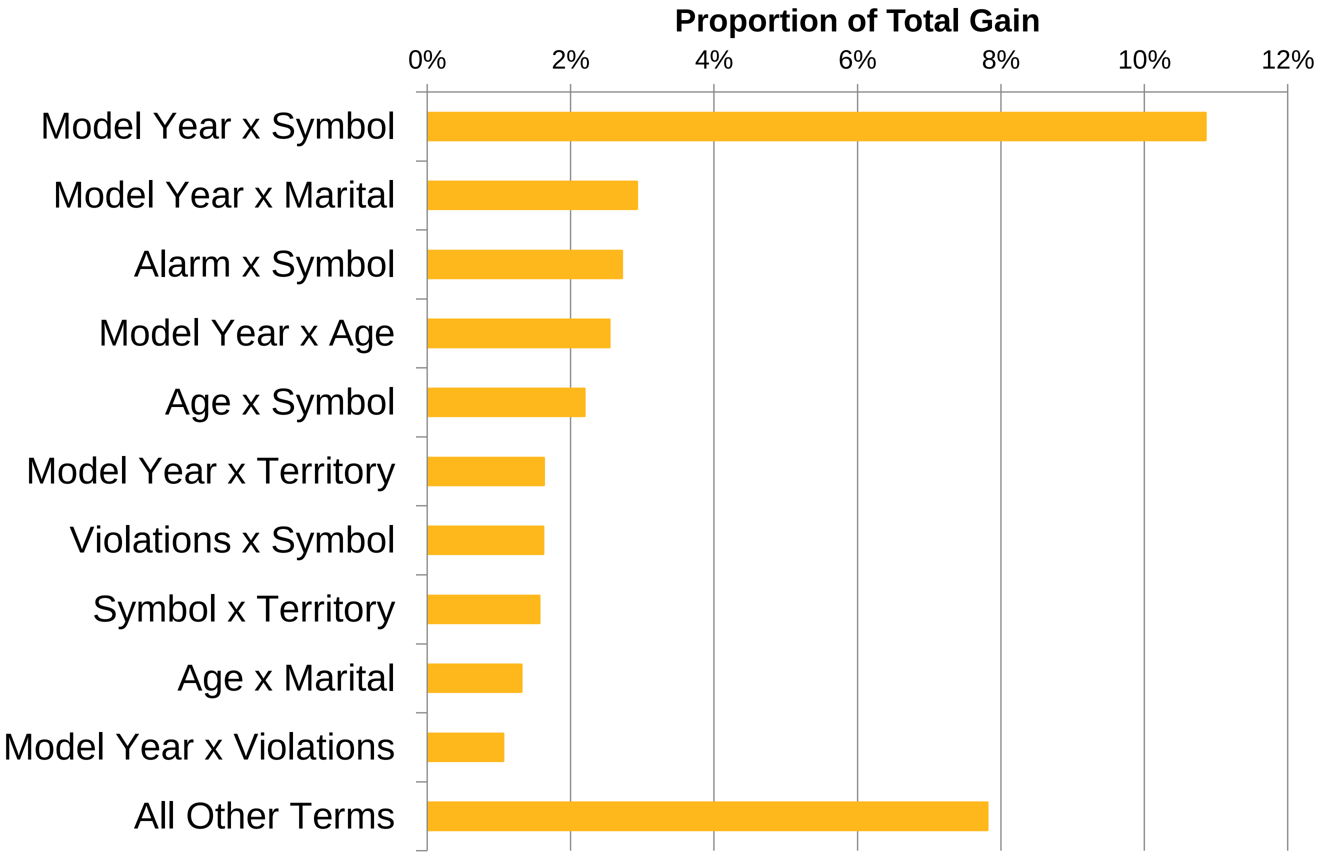


Use Case 1: Selecting factors to review

Proportion of Total Gain

Variable	Factor Importance
Model Year x Symbol	11%
Model Year x Marital	3%
Alarm x Symbol	3%
Model Year x Age	3%
Age x Symbol	2%
Model Year x Territory	2%
Violations x Symbol	2%
Symbol x Territory	2%
Age x Marital	1%
Model Year x Violations	1%
All Other Terms	8%

Factor Importance

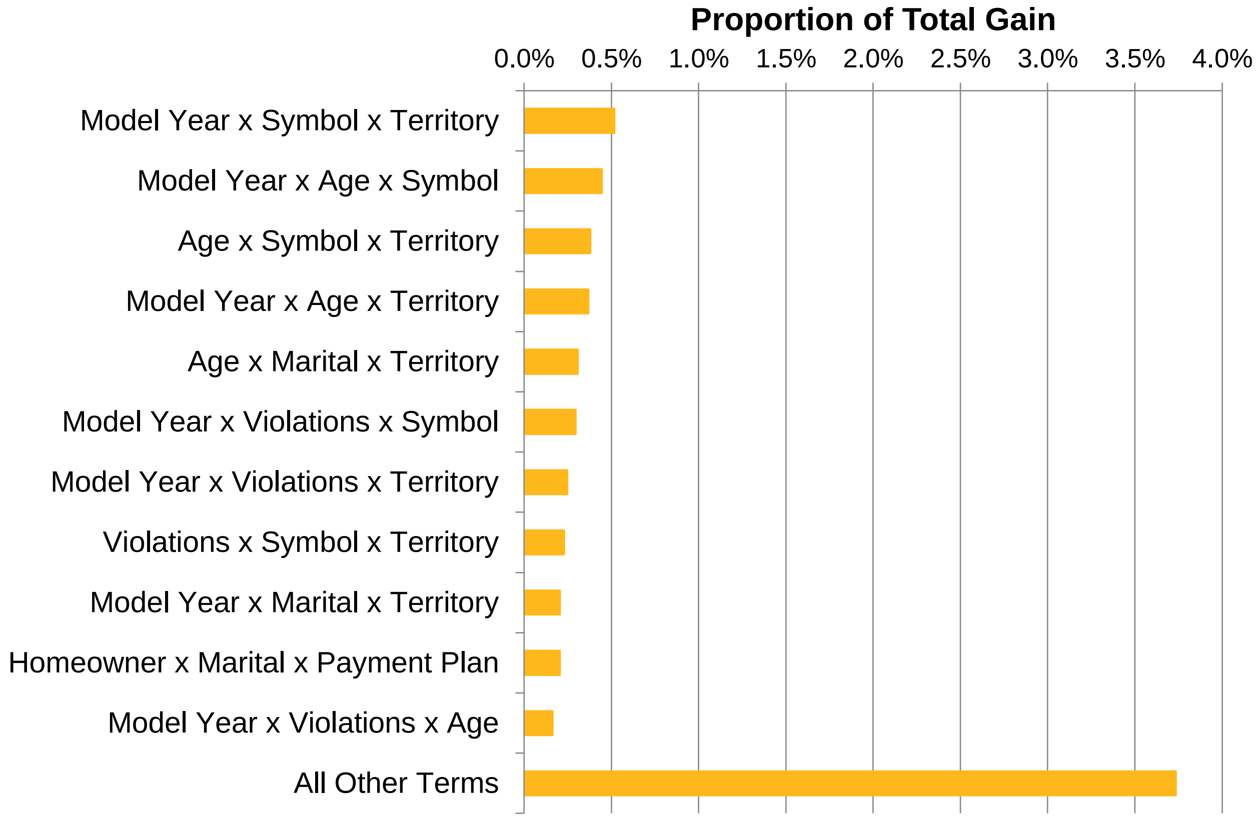


Use Case 1: Selecting factors to review

Proportion of Total Gain

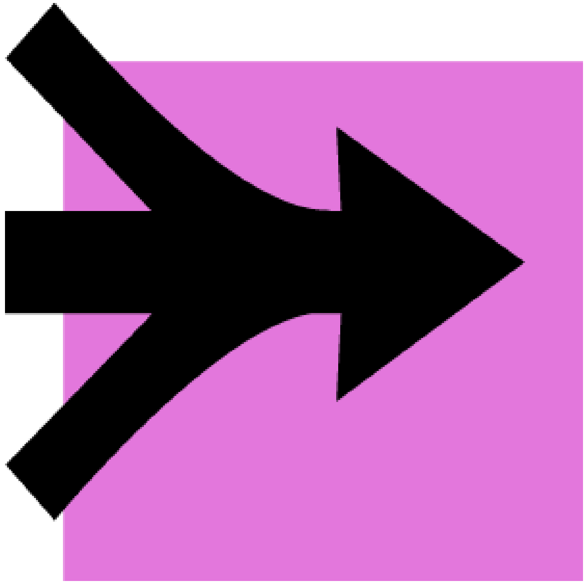
Variable	Factor Importance
Model Year x Symbol x Territory	0.5%
Model Year x Age x Symbol	0.4%
Age x Symbol x Territory	0.4%
Model Year x Age x Territory	0.4%
Age x Marital x Territory	0.3%
Model Year x Violations x Symbol	0.3%
Model Year x Violations x Territory	0.2%
Violations x Symbol x Territory	0.2%
Model Year x Marital x Territory	0.2%
Homeowner x Marital x Payment Plan	0.2%
Model Year x Violations x Age	0.2%
All Other Terms	3.7%

Factor Importance



Use Case 2: Find interactions to include in your GLM

- You are building a model for underwriting to determine which of your independent agents should have their book of business reviewed.
- The following variables are available for analysis:



Variables

- Agent Tenure
- % of Writeoff
- Exposures
- Affinity
- Avg Veh Age
- % with Convictions
- Multiple line
- Annual PMT

Use Case 2: Find interactions to include in your GLM

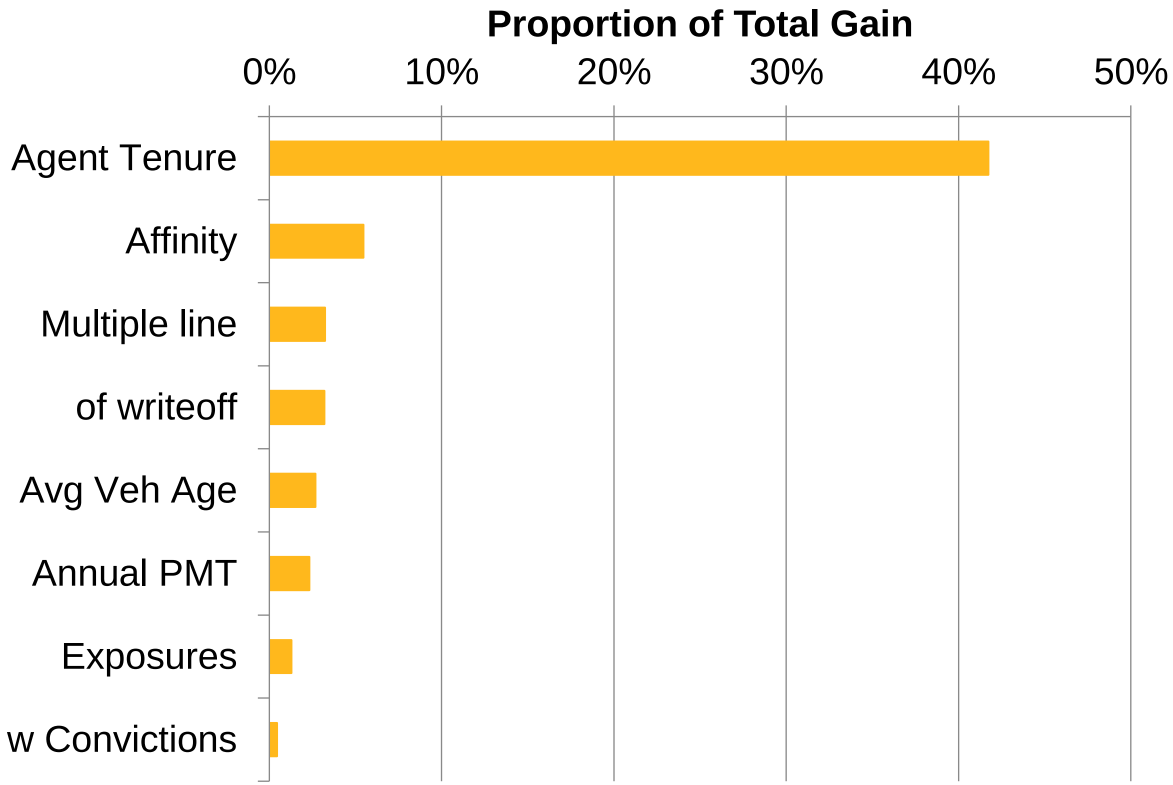
Variable	Exp(Value)	SE %
Agent Tenure	0.76	29.6
Affinity	1.42	22.9
Multiple line	NA	NA
% of writeoff	1.11	189.3
Avg Veh Age	NA	NA
Annual PMT	1.83	50.1
Exposures	NA	NA
% w Convictions	NA	NA

28 potential interactions

Use Case 2: Find interactions to include in your GLM

Proportion of Total Gain	
Variable	**Factor Importance**
Agent Tenure	42%
Affinity	5%
Multiple line	3%
% of writeoff	3%
Avg Veh Age	3%
Annual PMT	2%
Exposures	1%
% w Convictions	0%

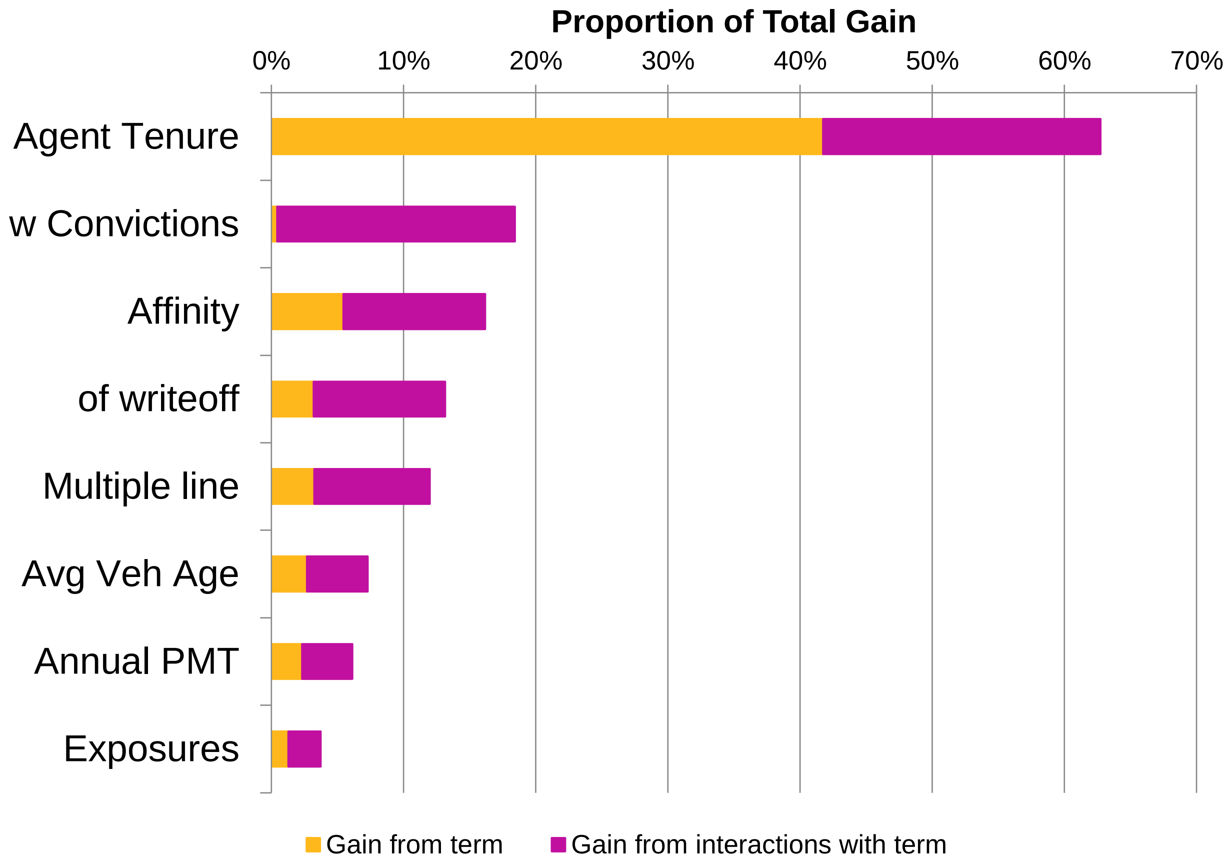
Factor Importance



Use Case 2: Find interactions to include in your GLM

Proportion of Total Gain		
Variable	**Gain from term**	**Gain from interactions with term**
Agent Tenure	42%	21%
% w Convictions	0%	18%
Affinity	5%	11%
% of writeoff	3%	10%
Multiple line	3%	9%
Avg Veh Age	3%	5%
Annual PMT	2%	4%
Exposures	1%	2%

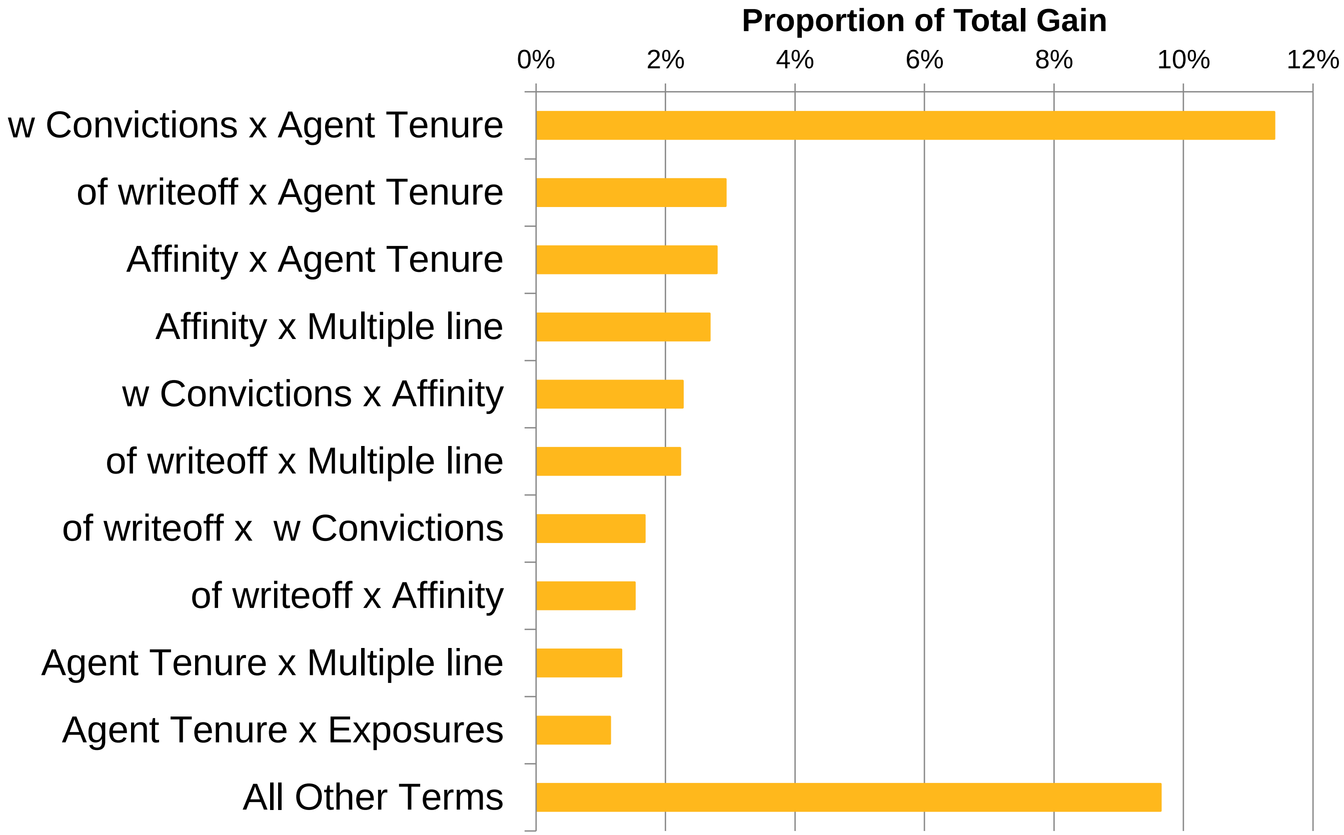
Factor Importance



Use Case 2: Find interactions to include in your GLM

Proportion of Total Gain	
Variable	**Factor Importance**
% w Convictions x Agent Tenure	11%
% of writeoff x Agent Tenure	3%
Affinity x Agent Tenure	3%
Affinity x Multiple line	3%
% w Convictions x Affinity	2%
% of writeoff x Multiple line	2%
% of writeoff x w Convictions	2%
% of writeoff x Affinity	2%
Agent Tenure x Multiple line	1%
Agent Tenure x Exposures	1%
All Other Terms	10%

Factor Importance

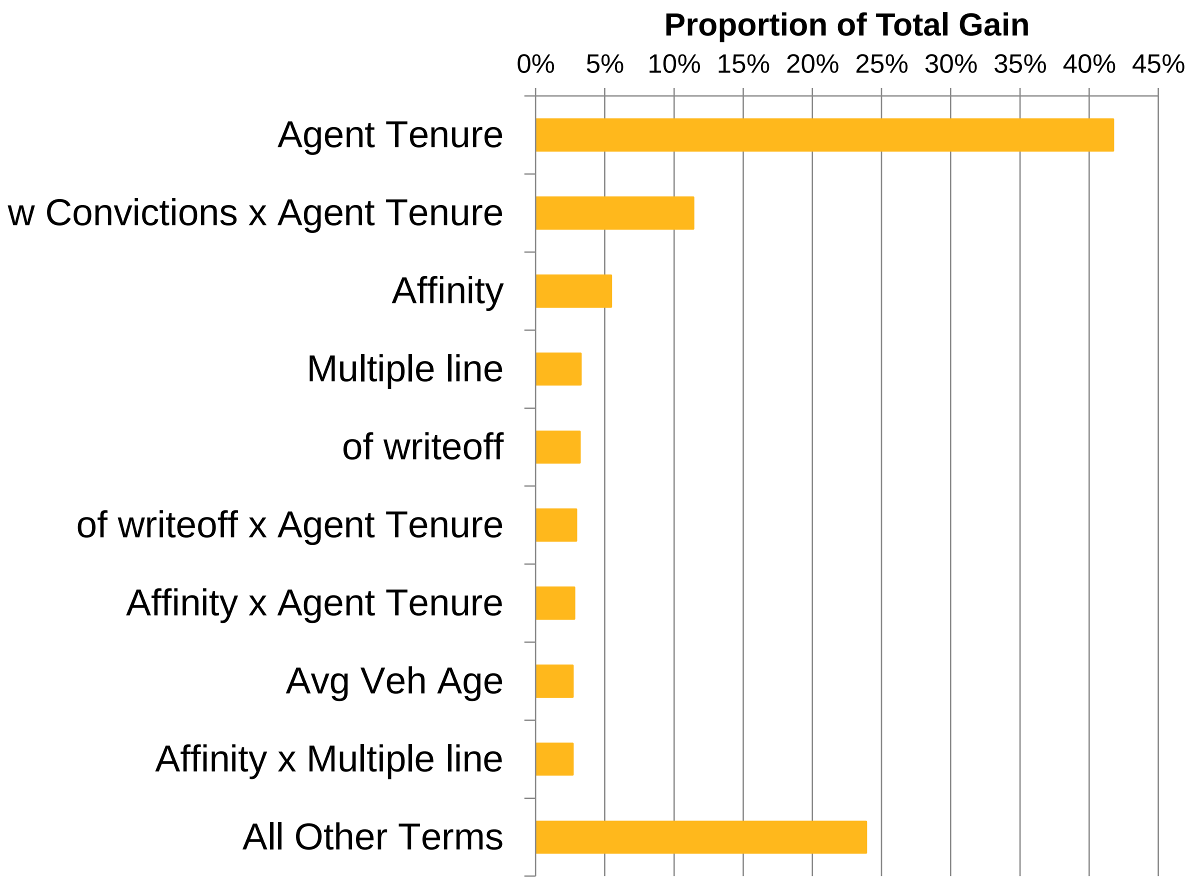


Use Case 2: Find interactions to include in your GLM

Proportion of Total Gain

Variable	Factor Importance
Agent Tenure	42%
% w Convictions x Agent Tenure	11%
Affinity	5%
Multiple line	3%
% of writeoff	3%
% of writeoff x Agent Tenure	3%
Affinity x Agent Tenure	3%
Avg Veh Age	3%
Affinity x Multiple line	3%
All Other Terms	24%

Factor Importance



Thank you