european actuarial academy

The European knowledge centre for actuaries

Combining GLMs with GBMs for the Best of Your Pricing Models

Piotr Lebiedź

ABOUT ME

Piotr Lebiedź

Guidewire
PricingCenter

- A seasoned pricing actuary, serving as Pricing Manager at Guidewire, driving adoption of a new pricing solution – Guidewire PricingCenter – across the globe.
- Specialized in empowering pricing leaders with the insights and tools necessary to navigate the ever-evolving pricing landscape.
- A passionate advocate for next-generation insurance pricing, committed to bridging actuarial expertise with cutting-edge technology to drive innovation across the industry.
- A university lecturer leading the actuarial path at Gdańsk University of Technology (Gdańsk, Poland).

Agenda

1. Quick intro to insurance pricing
2. GLM vs. GBM – comparison
   i. definition and concept
   ii. strengths and weaknesses
   iii. typical applications in insurance pricing
3. GLM & GBM – hybrid approach
   i. GBM helping in variable selection for GLM
   ii. GBM as input to GLM
   iii. GLM as input to GBM
   iv. GLM residuals corrected by GBM
   v. GLM and GBM ensembled
4. Key takeaways

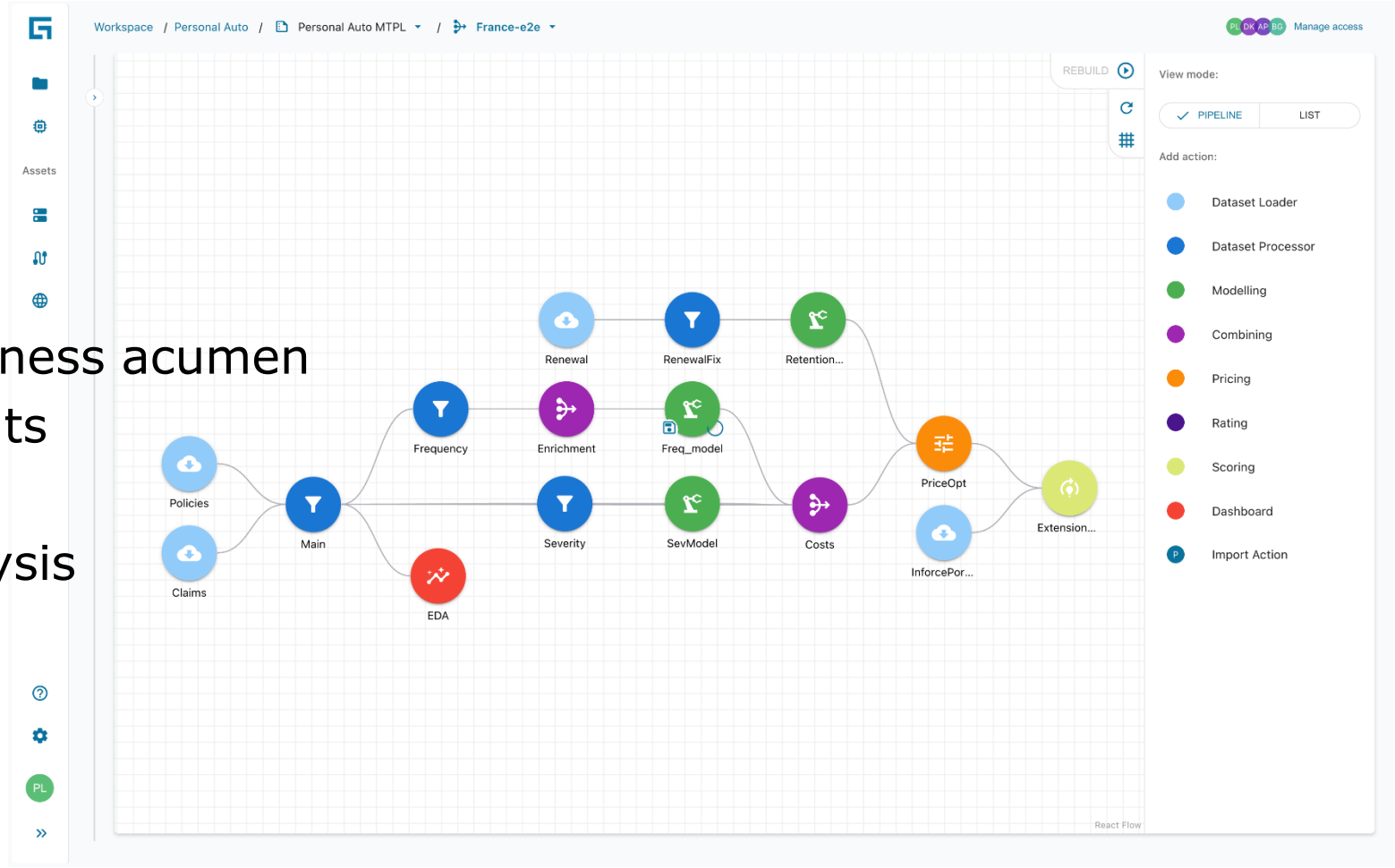
Agenda

1. **Quick intro to insurance pricing**
2. GLM vs. GBM – comparison
   i. definition and concept
   ii. strengths and weaknesses
   iii. typical applications in insurance pricing
3. GLM & GBM – hybrid approach
   i. GBM helping in variable selection for GLM
   ii. GBM as input to GLM
   iii. GLM as input to GBM
   iv. GLM residuals corrected by GBM
   v. GLM and GBM ensembled
4. Key takeaways

Insurance pricing aims to calculate fair price for particular risk transfer

Insurance pricing consists of:

- Data wrangling
- Data exploration
- Product understanding
- Risk modeling
- Expert judgement and business acumen
- Forward-looking adjustments
- Market analysis
- Competitor landscape analysis
- Price optimization
- Frequent price updates
- Constant monitoring
- And more…



Agenda

1. Quick intro to insurance pricing
2. **GLM vs. GBM – comparison**
   i. definition and concept
   ii. strengths and weaknesses
   iii. typical applications in insurance pricing
3. GLM & GBM – hybrid approach
   i. GBM helping in variable selection for GLM
   ii. GBM as input to GLM
   iii. GLM as input to GBM
   iv. GLM residuals corrected by GBM
   v. GLM and GBM ensembled
4. Key takeaways

Definition and concept

GLM

- Full name: Generalized Linear Model
- Category: Linear estimator
- Domain: Identifying, quantifying, and combining independent univariate signals

GLM formula:

$$g\big(E(Y)\big) = X\beta$$

where:
g is a link function
$E(Y)$ is the expected value of the response variable Y
X is a matrix of predictor variables,
β is a vector of coefficients

GBM

- Full name: Gradient Boosting Machine
- Category: Tree-based estimator
- Domain: Identifying multivariate dependencies and microsegments
- Ensemble learning method: Boosting

GBM builds decision trees iteratively and corrects their combined errors on the way. Thanks to a gradient descent algorithm it minimizes the loss when adding new models.

Strengths and weaknesses

GLM

+ Well-known and market standard
+ Transparent formula once trained
+ Easily ingested by rating engines
+ Quick in real-time quoting

- Limited accuracy
- Assumes linearity
- Assumes feature independence
- Long time to build
- Subject to confirmation bias

GBM

+ Best-in-class accuracy
+ Natively catches non-linearity
+ Finds interactions and microsegments
+ Quick setup, little feature transformations

- Less-known and less-practiced
- Lack of transparency (black-box)
- Requires advanced engine for prod use
- Significant latency in real-time quoting
- Requires more data and is easier to overfit

Typical applications in insurance pricing

GLM

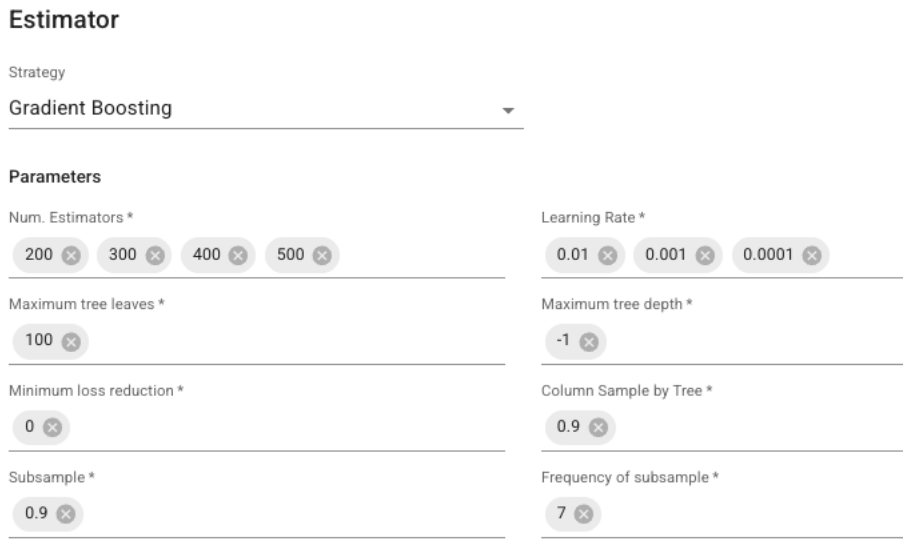
1. Risk models
2. Demand/elasticity models
3. Other models (churn, market price)

GBM

1. Residual risk models
2. Market price models
3. Other models (risk, demand, churn)

Feature	Coefficient	Relativity	Std. Error	p-value	Significance
INTERCEPT	-3.127	0.044	0.054	<.001	***
Features					
BONUS_MALUS					
BONUS_MALUS.BIN=[-inf;51.5)	0	1			
BONUS_MALUS.BIN=[51.5;68.5)	0.619	1.858	0.019	<.001	***
BONUS_MALUS.BIN=[68.5;72.5)	0.687	1.987	0.039	<.001	***
BONUS_MALUS.BIN=[72.5;76.5)	0.739	2.093	0.040	<.001	***
BONUS_MALUS.BIN=[76.5;85.5)	0.851	2.342	0.030	<.001	***
BONUS_MALUS.BIN=[85.5;95.5)	0.895	2.447	0.032	<.001	***
BONUS_MALUS.BIN=[95.5;inf)		5.160	0.028	<.001	***
DENSITY					

BONUS_MALUS.BIN=[85.5;95.5)

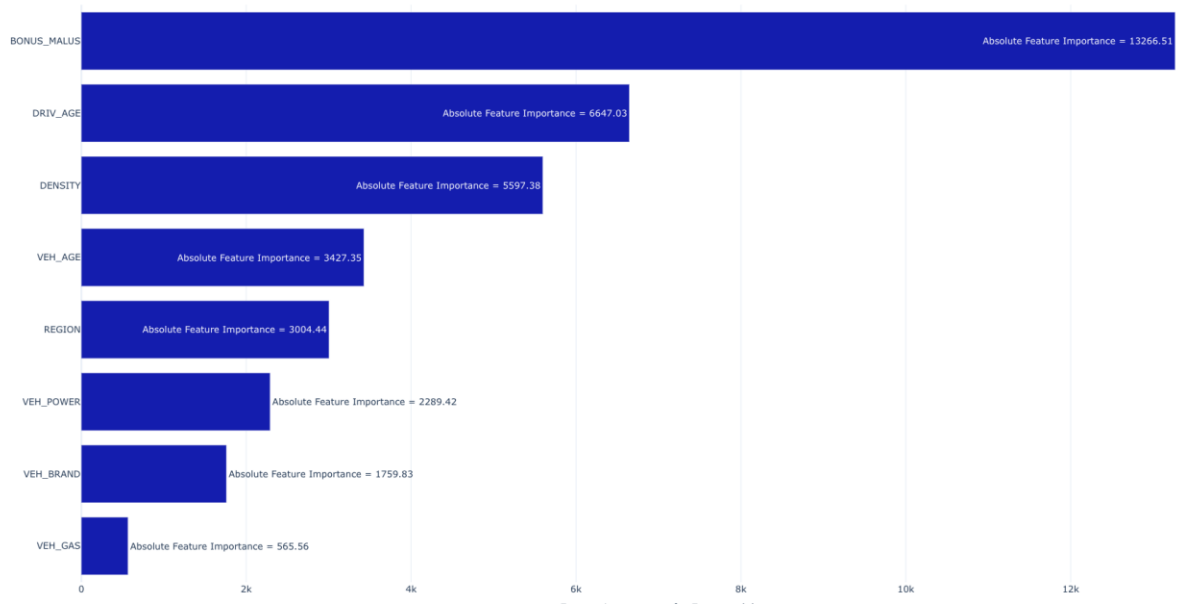


Agenda

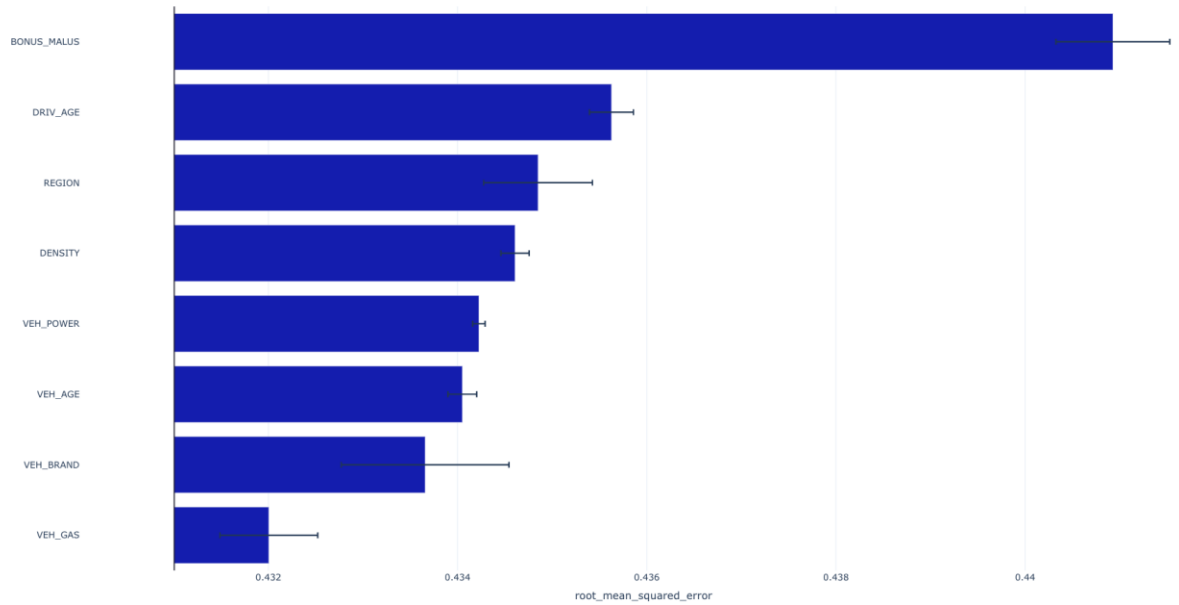
1. Quick intro to insurance pricing
2. GLM vs. GBM – comparison
   i. definition and concept
   ii. strengths and weaknesses
   iii. typical applications in insurance pricing
3. **GLM & GBM – hybrid approach**
   i. GBM helping in variable selection for GLM
   ii. GBM as input to GLM
   iii. GLM as input to GBM
   iv. GLM residuals corrected by GBM
   v. GLM and GBM ensembled
4. Key takeaways

GBM helping in variable selection for GLM

Tree-based feature importance

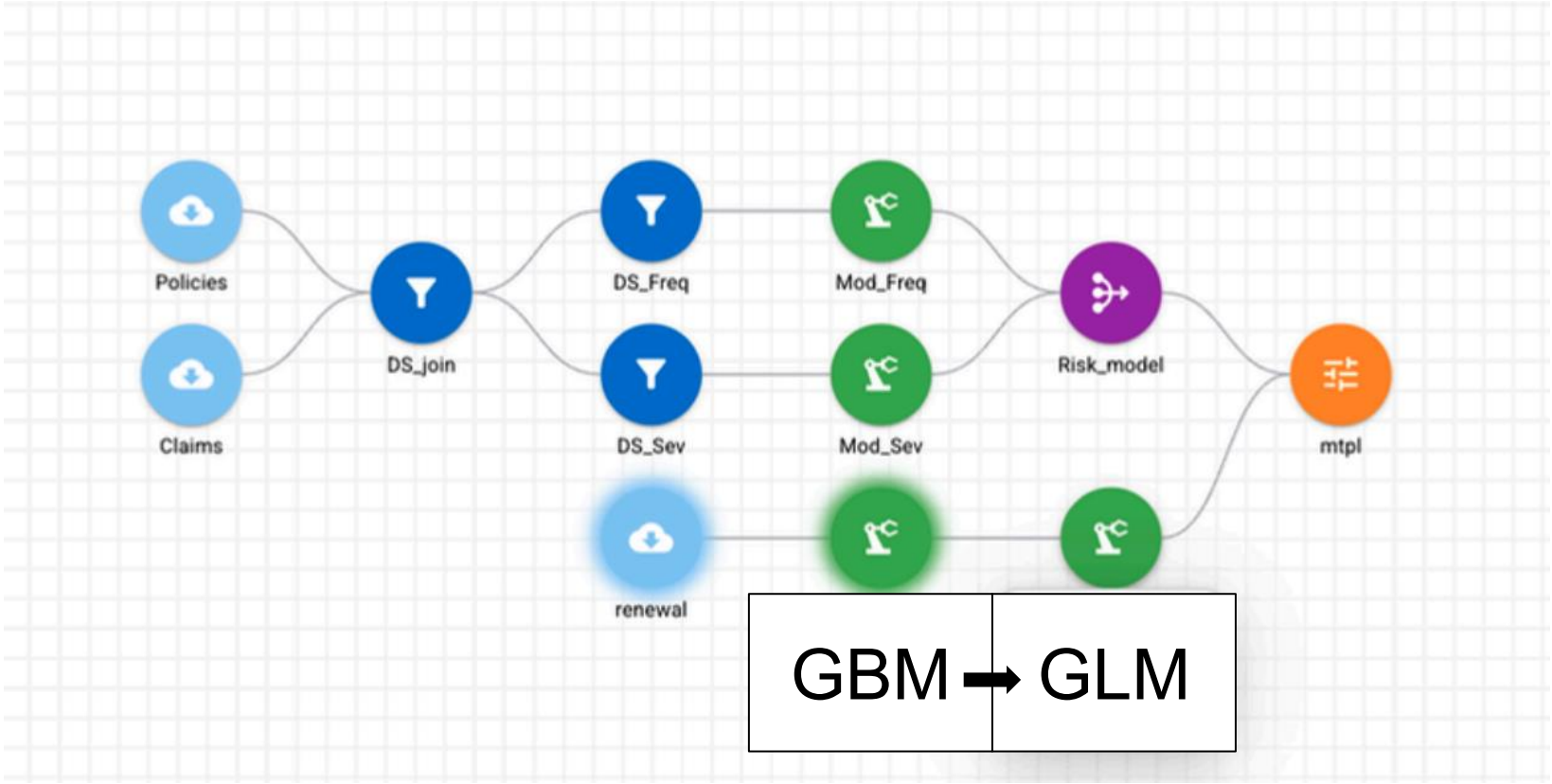


Model agnostic variable importance



GBM as input to GLM

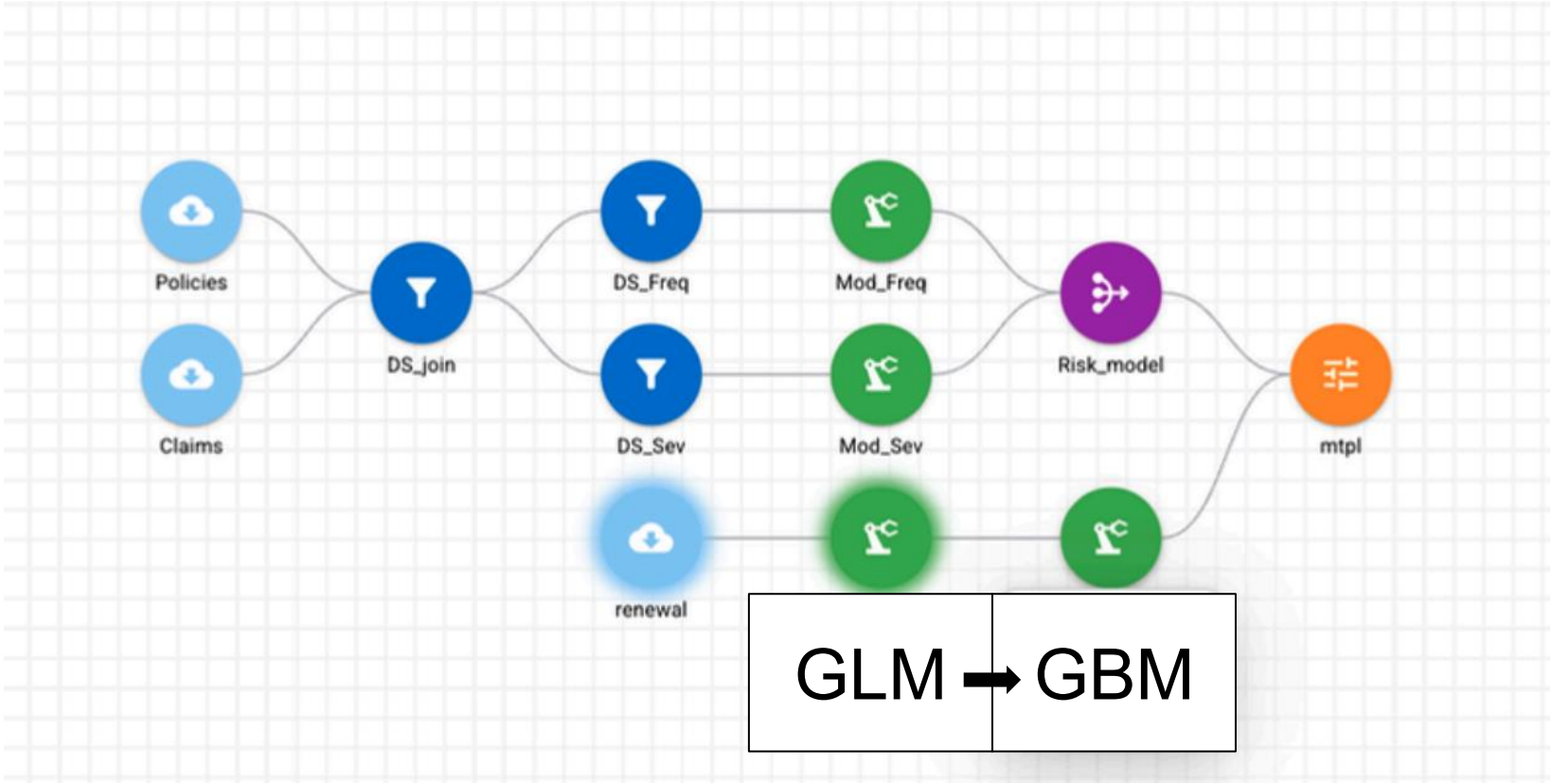
- Prediction = GLM(GBM, x1, x2, x3, ...)
- Controlled impact of GBM



GLM & GBM – hybrid approach

GLM as input to GBM

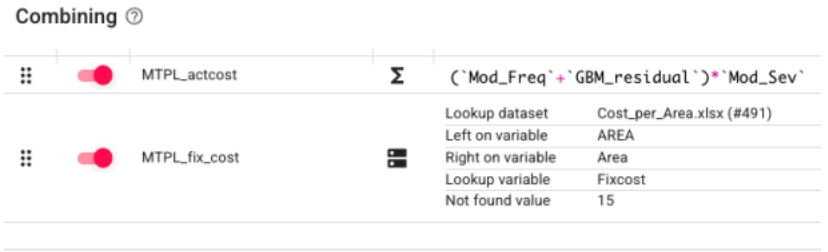
- GBM's behavior controlled by adding predictions from GLM as input
- Allowing GBM to find more nuances, keeping the original target distribution



GLM & GBM – hybrid approach

GLM residuals corrected by GBM

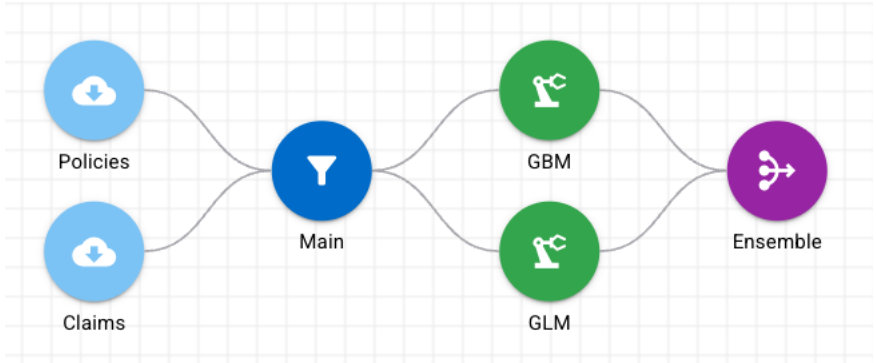
1. Fit GLM and get predictions y_{GLM}
2. Compute residuals: $actual\ value\ -\ prediction$
3. Train GBM on residuals (loss matching distribution of residuals) and get predictions $\bar{y}_{GBM}$
4. Update predictions: $y_{GLM} + \bar{y}_{GBM}$

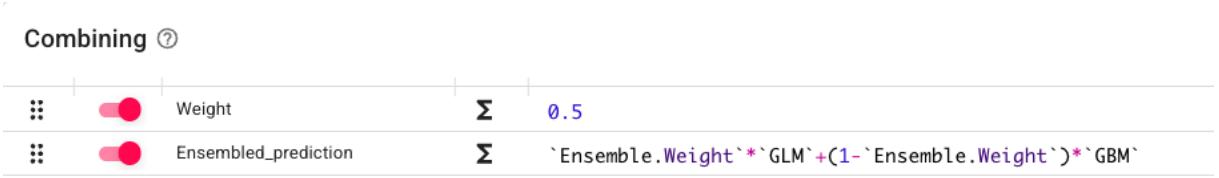


GLM & GBM – hybrid approach

GLM and GBM ensembled

1. Fit GLM and get predictions y_{GLM}
2. Fit GBM and get predictions y_{GBM}
3. Choose weights *w* between 0 and 1
4. Final predictions: $w \cdot y_{GLM} + (1-w) \cdot y_{GBM}$





Agenda

1. Quick intro to insurance pricing
2. GLM vs. GBM – comparison
   i. definition and concept
   ii. strengths and weaknesses
   iii. typical applications in insurance pricing
3. GLM & GBM – hybrid approach
   i. GBM helping in variable selection for GLM
   ii. GBM as input to GLM
   iii. GLM as input to GBM
   iv. GLM residuals corrected by GBM
   v. GLM and GBM ensembled
4. **Key takeaways**

To summarise

- GLM and GBM are different tools in the actuarial toolkit

- GLM assesses individual signals well and is fully transparent
- GBM easily finds multivariate dependencies and microsegments but is 'black-box'
- GLM requires more actuarial expertise, while GBM requires data science rigour

- There is no 'better' or 'worse' between them
- They have different advantages and may serve different purposes
- One person or team can use both, either separately – for different cases, or…
- … combine them for the same task to leverage the best of both

- Best-in-class pricing solutions are capable of building both kinds of models, applying a hybrid approach, and deploying them seamlessly to production

european actuarial academy

The European knowledge centre for actuaries

Thank you!

Please rate the conference via the survey-link you will receive per email.

Visit our website

www.actuarial-academy.com

for more events.

Follow us on LinkedIn

www.linkedin.com/company/642904

for updates & actuarial fun.