Akur8 GLM: Methodology and Regulatory review

May 24, 2022

Agenda

Akur8 provides software for the specific use case of ratemaking

Introduction to Akur8 GLM

- Advantages of Penalized GLMs
- Differences and similarities to traditional GLMs
- Blending Credibility and GLMs
- Introduction to Smoothness

Akur8 GLM in Practice

- Modeling Flow
- Smoothness in Practice
- Methods of Evaluation
- Cross Validation

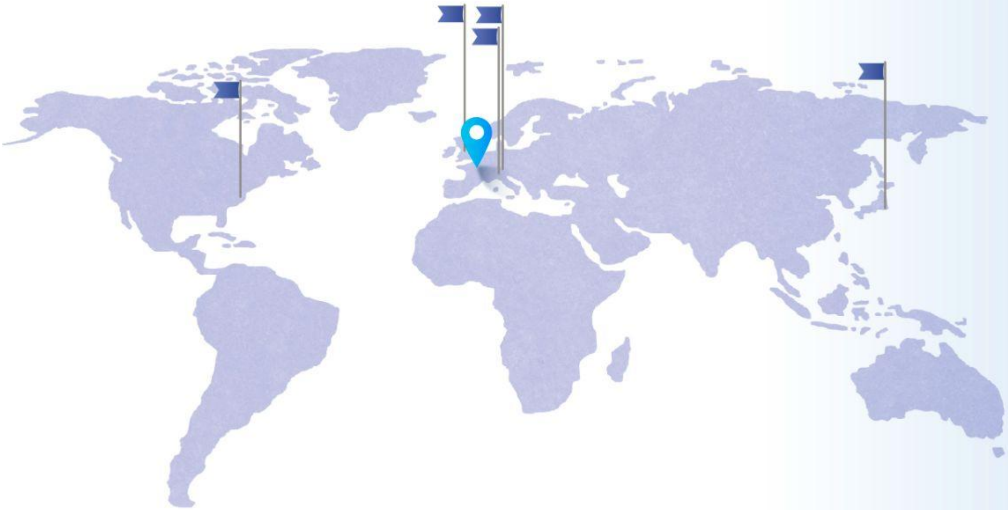
Regulatory Review

- Filing support template
- Exhibit Templates
  - Goodness of Fit
  - Variable Plots
  - Spread
- Conclusions

Introduction to Akur8 GLM

AKUR8

Company overview



Founded
2018

Global offices
NYC, Paris, London, Munich, Milan, Tokyo

Employees
100

Nationalities
20+

Activity
Non-Life Insurance Pricing (e.g. P&C, Health, Travel)

Customers
50+

The software

Akur8 provides software for the specific use case of ratemaking

Akur8 provides a cloud based **software** to actuaries and ratemaking modelers to build auditable models with efficiency and transparency.

Akur8 GLM methodology

- Outputs the factors in a **traditional table based format**;
- Combines the traditionally accepted **actuarial methodologies of GLMs and Credibility assumptions**;
- Is the result of **7 years of R&D** and is **a next generation ratemaking tool.**

Filing Support Documentation

- **Exhibits Templates** are generated by Akur8 software;
- Follows the **best practices** suggested in Predictive Model Whitepaper and Speed to Market presentation by **NAIC**;
- Promotes **rigor and standardization** in filing that benefits all stakeholders, including regulators, carriers, and other participants;
- Was designed and developed in collaboration with **Milliman**.

Akur8 GLM

A Penalized Regression or Regularization Method

"Akur8 GLM" is a modeling technique that blends standard GLMs with additional constraints, making it a **Penalized Regression.** Penalized Regressions are standard in modeling practices, being studied and published for more than 20 years (see references).

They are becoming increasingly popular in insurance applications:

Lasso, Ridge and Elastic net (Glmnet)

- Presented on the NAIC 2021 June Book club - Regularization Method.
- Section 10.5 in the CAS Monograph - Generalized Linear Models for Insurance, Rating Second Edition

Advantages of Penalized GLMs

Blending Credibility and incorporating non-linearities

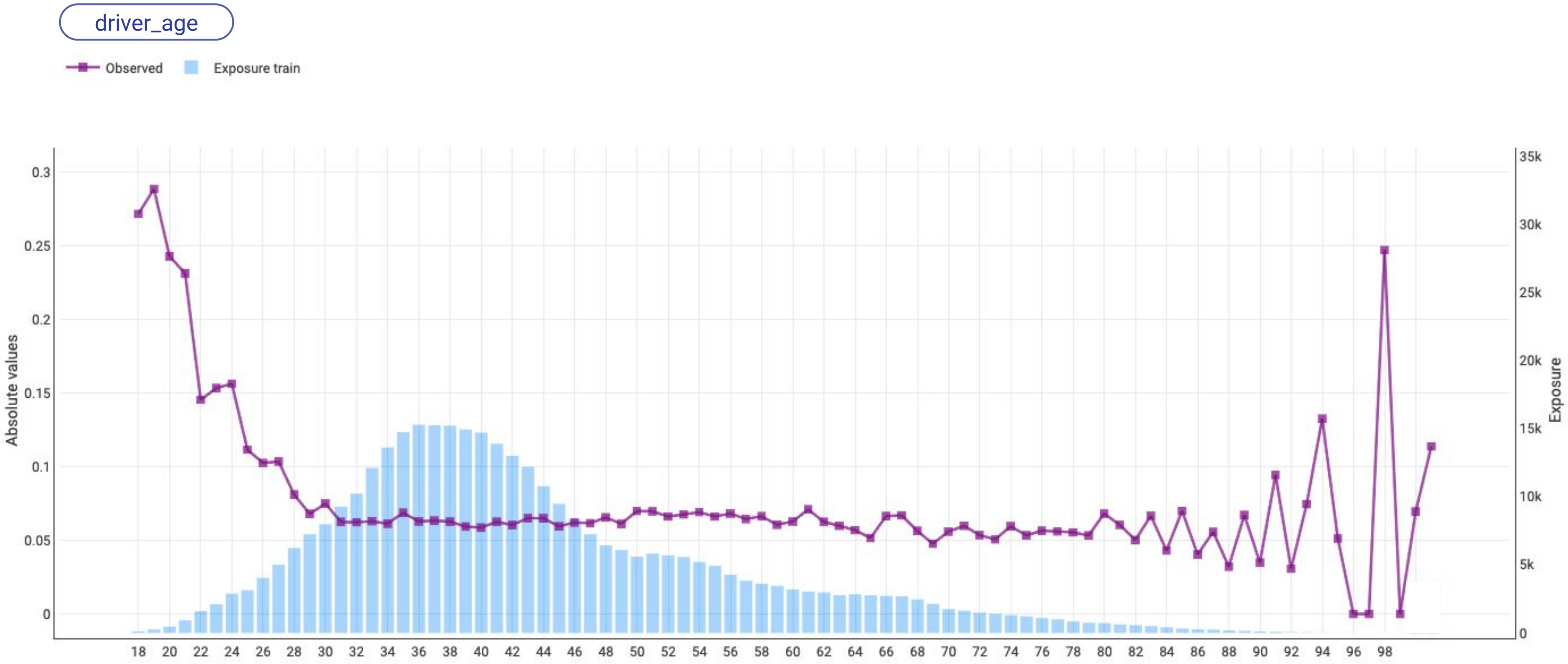
Blends Credibility with a GLM

- In a segment with little data, standard GLMs do not provide accurate estimates of the risk, as they tend to overfit;
- Penalized GLMs avoid overfitting by treating small segments as if they are not fully credible.

Natively fits non linear effects

- Non linearities are often required to provide sound estimations of the risk;
- Feature engineering addresses that but can be arbitrary and can lead to improper estimations in segments with low data.

Advantages: illustrative example



Advantages: illustrative example

NON LINEARITY

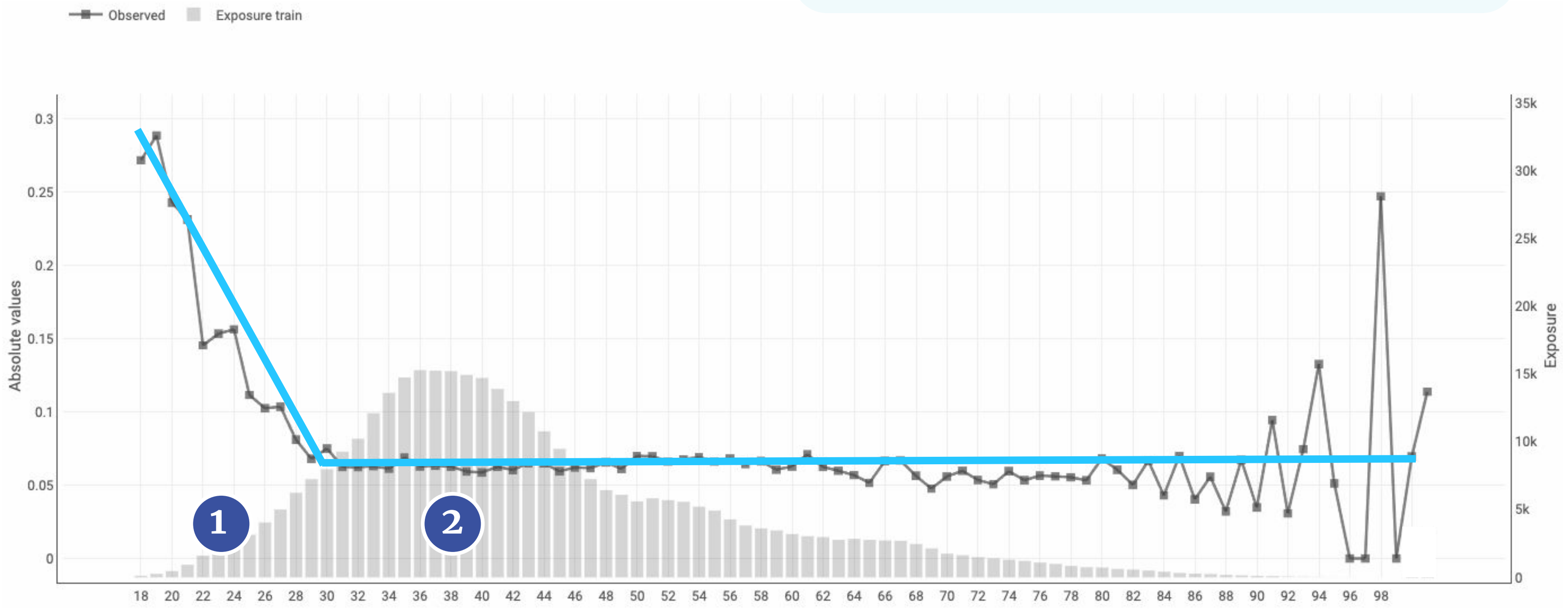
The shape for the driver_age variable is non-linear. It intuitively can be split in 2 different behaviours.

Advantages: illustrative example

NON LINEARITY

The shape for the driver_age variable is non-linear. It intuitively can be split in 2 different behaviours.

driver_age

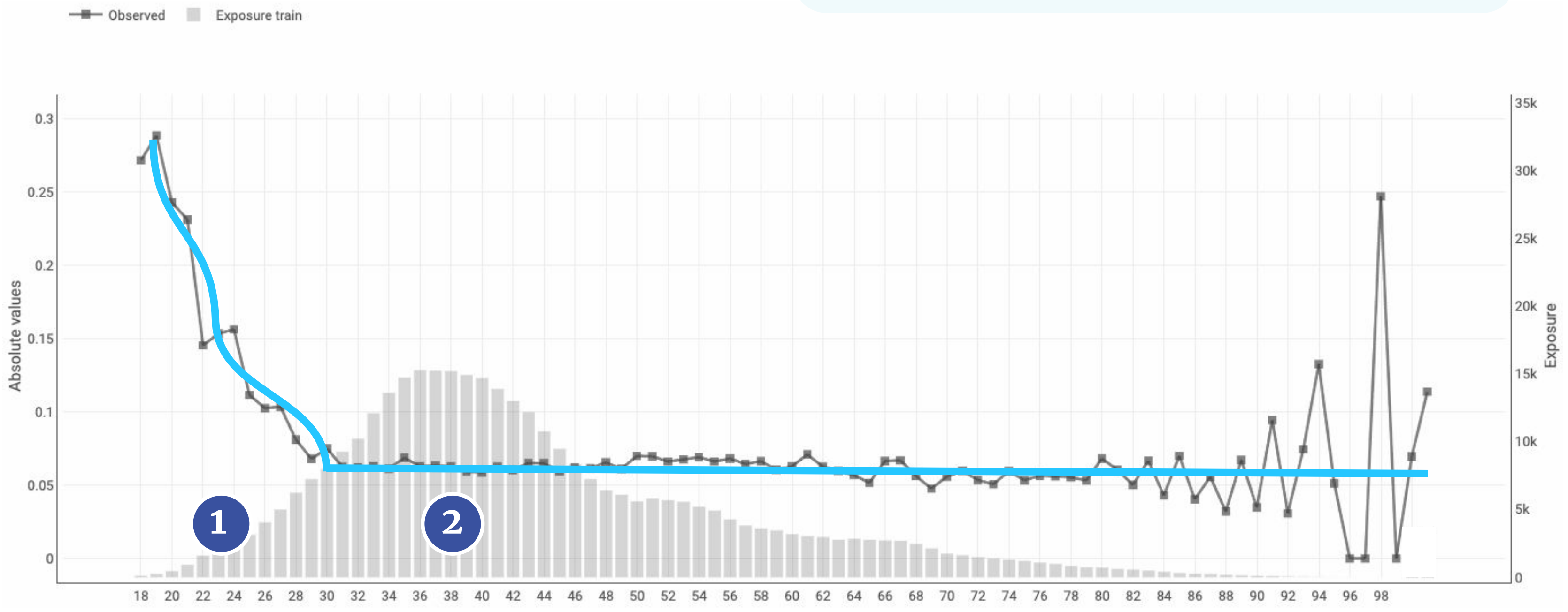


Advantages: illustrative example

NON LINEARITY

The shape for the driver_age variable is non-linear. It intuitively can be split in 2 different behaviours.

driver_age

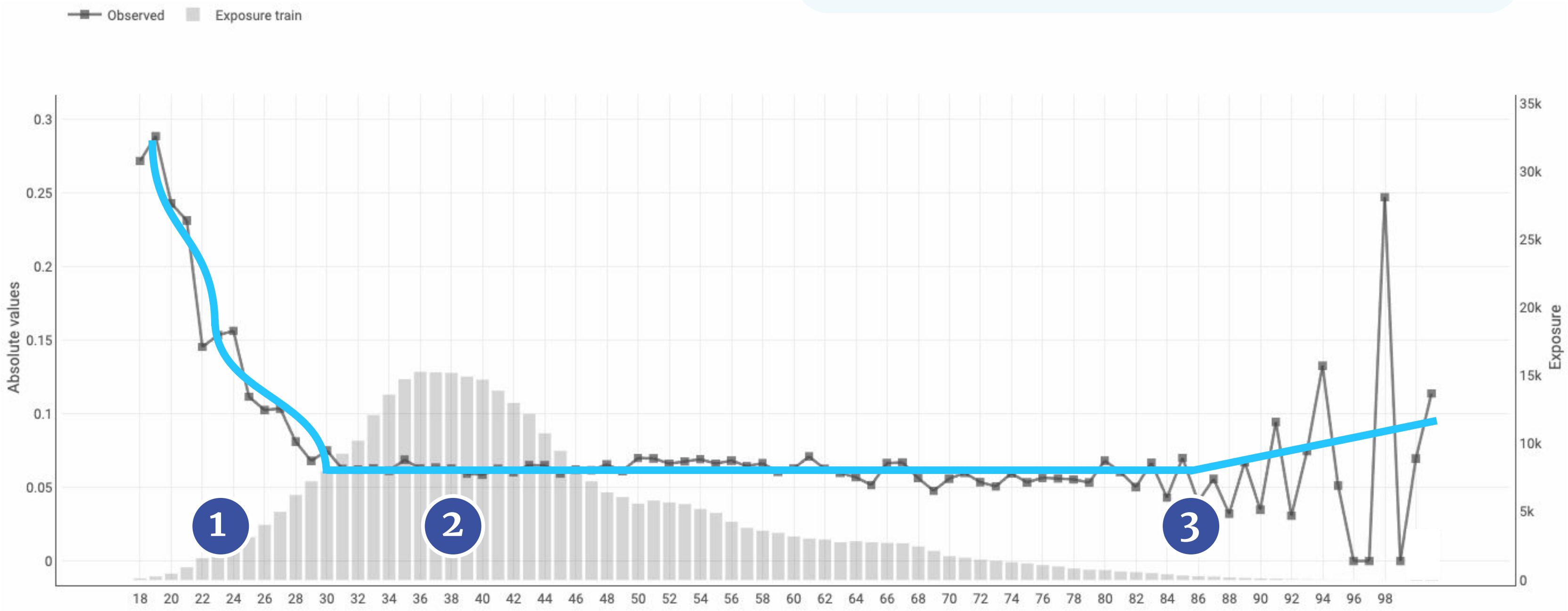


Advantages: illustrative example

NON LINEARITY

The shape for the driver_age variable is non-linear.
It intuitively may be split in 3 different behaviours.

driver_age

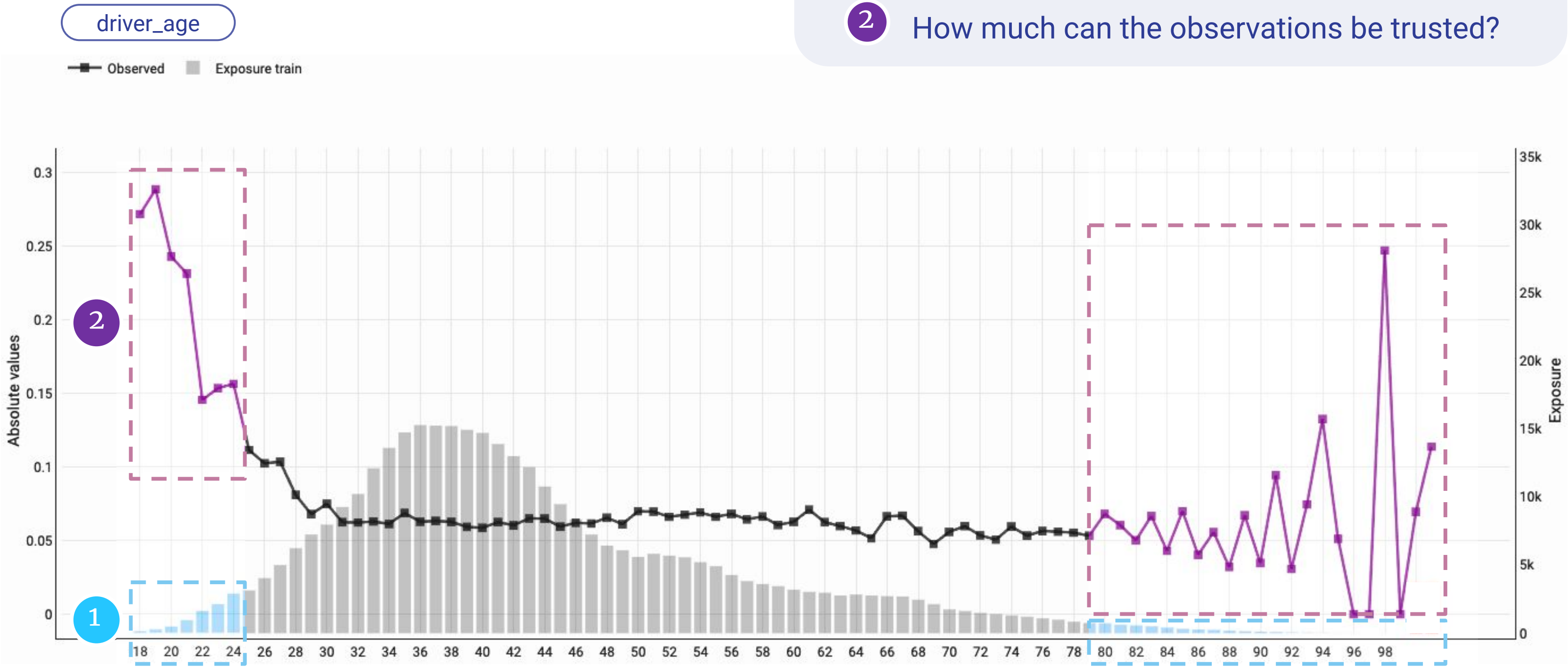


Advantages: illustrative example

CREDIBILITY

1 Young and old segment have limited data

2 How much can the observations be trusted?



Similarities with standard GLMs

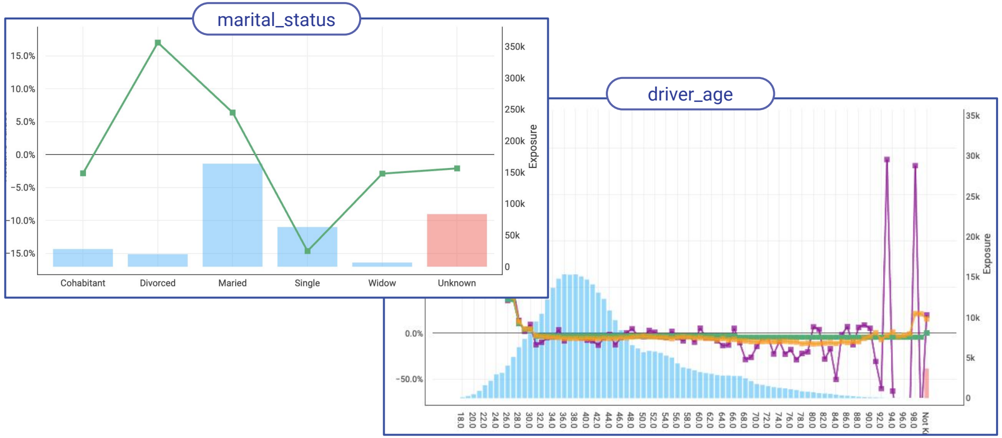
Table based models with standard distribution and assumptions

Akur8 GLMs are standard GLMs with additional constraints on the coefficients.

Similar to standard GLMs, models created with Akur8 methodology:

- Have **standard statistical assumptions** - observations follow commonly used distributions and link functions;
- Produce **table based outputs**: the model's predictions are calculated by multiplication of coefficients [1] associated with each variable.

Variable	Level	Indicated
Base	Base	0.062
marital_status	Divorced	1.170
marital_status	Married	1.064
marital_status	Single	0.854
marital_status	Widow	0.971
marital_status	Unknown	0.979
driver_age	16	4.209
driver_age	17	4.209
driver_age	18	2.422
driver_age	19	2.422
driver_age	20	2.422
driver_age	21	2.422
driver_age	22	2.072
driver_age	23	1.968



[1] Multiplications are used for logarithmic link models; in general, GLMs predictions are given by the inverse link of the sum of the linear predictors.

Differences with standard GLM

Maximizing the tradeoff between likelihood and penalty

GLMs are fit by maximizing the likelihood of a specific function via the formula below.

Penalized GLMs modify the GLM formula by adding an extra term, the Penalty.

The Penalty incorporates extra assumptions on the model, allowing to **blend Credibility with a GLM**:

- **Standard GLM:**

  Model = max (LogLikelihood(Observed))

- **Akur8 GLM and Penalized GLM:**

  Model = max (LogLikelihood(Observed) **– Penalty(Coefficient))**

Quick Reminder… What is Credibility

"Credibility, simply put, is the weighting together of different estimates to come up with a combined estimate."

Foundations of Casualty Actuarial Science

When the volume of data is not enough to accurately estimate the losses, Credibility methodologies provide ways to **complement the observed experience with additional information**.

The Credibility formula is:

Estimate = Z * **Observed Experience** + (1 – Z) * Related Experience

where the Credibility factor **Z** is a number between 0 and 1.

GLM and Credibility

"GLMs effectively assume that the underlying datasets are 100% credible, no matter their size.

If some segments have little data, the resulting uncertainty would not be reflected in the GLM parameter estimates themselves."

Regulatory Review of Predictive Model

When GLMs are used to estimate losses, the GLM formula becomes

GLM Estimate = **Observed Experience**

This implies that, for each category, a GLM will give 100% Credibility to the data regardless of underlying exposure.

In the equation below, the Likelihood is maximized when observations and predictions exactly match.

GLM Model = max(LogLikelihood(Observed))

Credibility: Worker’s Compensation example

Loss Cost by class code example

Losses and exposures for companies are collected, and we want to compute an estimation of the average loss cost per class code.

Credibility aims to find an appropriate **tradeoff** between average loss by class called Observed Experience (**purple lines**) or grand average called Complement of Credibility (**black dotted line**).

The plot represents the **observed losses**:

- The **blue bars** represent the number of observations for a given class;
- The **purple lines** represent the **Observed Experience** as the average loss cost for each class;
- The **black line** represent the **Complement of Credibility** as the overall average (or grand average) of $500 in this example.

Credibility: GLMs estimates

GLMs give 100% Credibility to the data

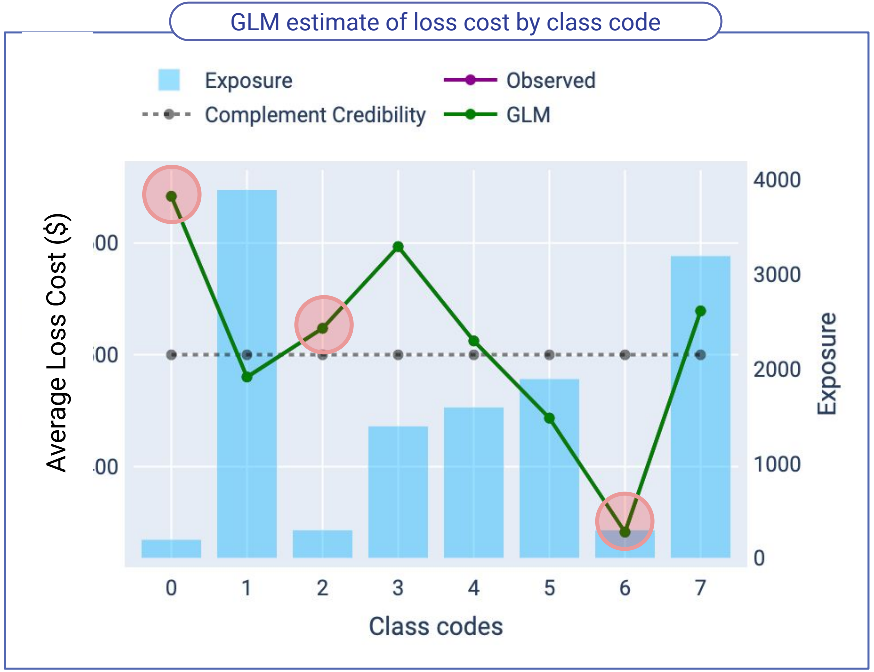
A GLM can be used to compute the estimates.

When the class code variable is modeled **categorically** [1] the estimates coincide with the observed data.

The **GLM estimate (green line)** exactly overlays the **Observed (purple line)**.

In "Credibility" terms, we are giving 100% Credibility to the data of every segment.

Such assumption **seems inappropriate** for class codes such as **0, 2,** and **6** which have low exposure.



[1] The design matrix for the GLM has one column for each of the class code, which has value 1 or 0 depending on whether the observation belongs to the relative class.

Credibility estimates

Credibility shrinks the estimates toward the grand average

The figure depicts the estimates when applying Credibility [1].

The Credibility estimates shrink toward the grand average (**black dotted line**) compared to observed and standard GLMs (**purple line**).

This is an outcome of the complement of Credibility being the grand average.

Furthermore, the lower the volume, the stronger the relative shrinkage.

[1] Buhlmann Credibility was used; the weight is computed by $Z = n / (n + k)$, with n the number of observations and k estimated via the data.

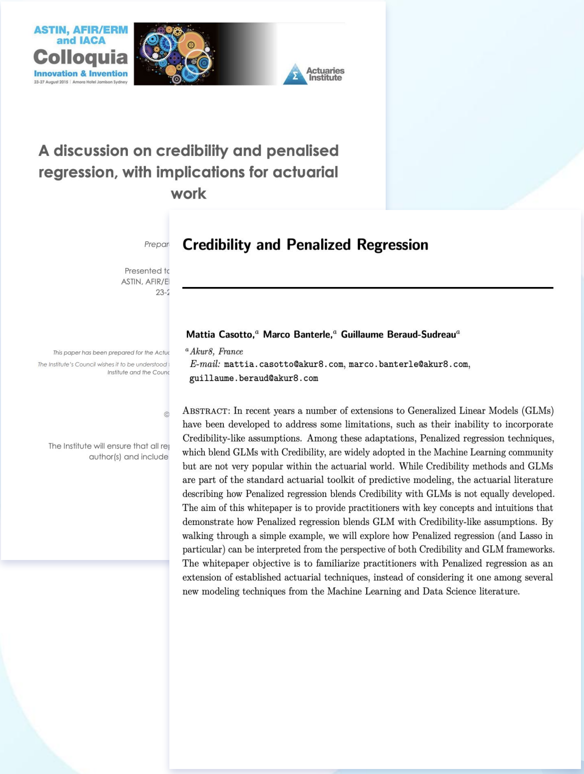
Blending GLM with Credibility

Penalized GLMs share the same properties as Credibility in the following ways:

1. **Shrink** GLM estimates toward the complement of Credibility (grand average);
2. Apply **more shrinkage** to segments with **low volume** of data / Credibility;
3. Equivalent to a **Bayesian model**, as in Buhlmann Credibility

   The theoretical connection between Credibility and Penalized GLM can be found in:

   - Fry, Taylor. "A discussion on credibility and penalised regression, with implications for actuarial work" (2015)
   - M.Casotto et al. "Credibility and Penalized Regression" (2022)



The GLM formula

The maximization formula that computes the classic GLMs estimates is given by:

$$\beta_{\mathrm{GLM}} = \max_{\beta} \mathrm{LogLikelihood}(\mathrm{Observed}, \beta)$$

- Makes the model replicate the training data;
- Builds models who give 100% Credibility to the data;
- Does not allow to incorporate prior knowledge on the model (Complement of Credibility).

The Penalized GLM Formula

Modifying the GLM maximization formula with extra terms introduces assumptions/priors on the coefficients.

$$\beta_{\text{Penalized}} = \max_{\beta} \text{LogLikelihood}(\text{Observed}, \beta) - \lambda_{\text{Smoothness}} \text{Penalty}(\beta)$$

The Likelihood

- Is the basic GLM formula.

The Smoothness

- Specifies the level of credibility to the data;
- When null, the Penalized model is equal to GLM.

The Penalty

- Represents the complement of credibility and the shrinkage.

The role of the penalty

The penalty encodes assumptions on the coefficients structure

In Akur8 GLM, the structure of the **penalty** encodes meaningful **assumptions/priors on the model**.

These assumptions **differ** on whether a variable is **categorical** or **ordinal** [1].

Categorical variable | Ordinal variable

The coefficients:

- **Are equal to zero** if they do not have enough evidence to deviate from the average;
- **Are more likely to be shrunk toward zero** if significantly different from the average.

[1] Ordinal variables are those where there exists a natural ordering of the levels. The design matrix is a 1/0 encoding equal to a categorical variable.

Example: **Smoothness and *Categorical* variables**

Raw data contains both **signal** and **noise.**

An appropriate model returns the good tradeoff between **robustness** and **sensitivity**.

A **robust** model gives more weight to the grand average (or complement of credibility) and requires a stronger signal to deviate from this level.

A **sensitive** model gives more weight to the relativities observed in the dataset and can deviate from the grand average more easily.

Example: **Smoothness and *Categorical* variables**

A **very strong smoothness** leads to models that are **extremely robust**.

The model captures the tendencies for the segments with highest exposure, keeping most coefficients are **equal to zero.**

The model **gives low credibility to the data** to most of the classes: for those class codes, the estimates are equal to the grand average.

Example: **Smoothness and *Categorical* variables**

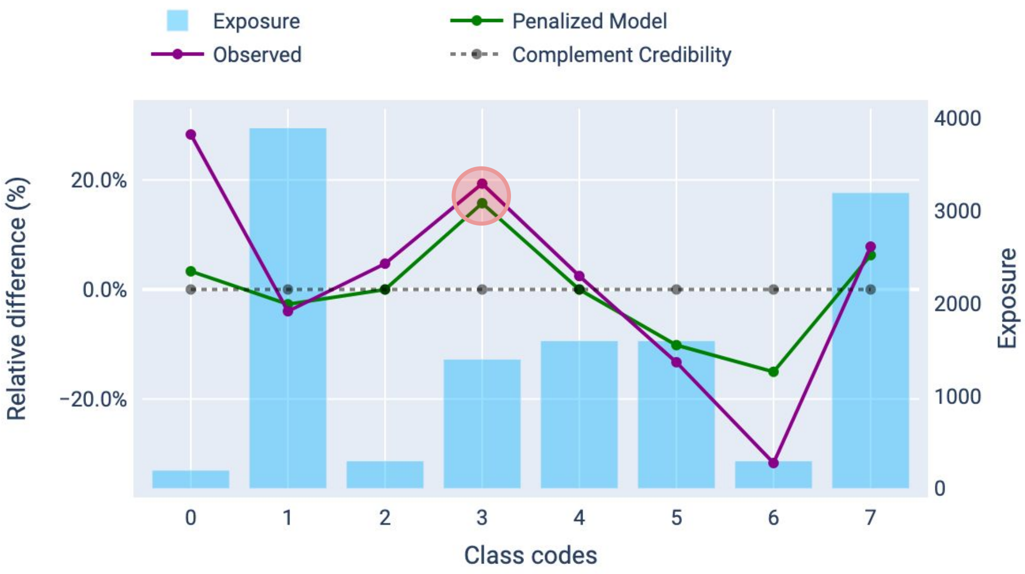
A very low value of the **smoothness** leads to models that are **overly sensitive**.

The output is **similar** to the output of a **GLM**: the estimate are close to the observations, regardless of whether the underlying volume of the data is credible enough.

Example: **Smoothness and *Categorical* variables**

A more balanced smoothness leads to **efficient models**.

The model offers a reasonable tradeoff between observed experience and complement of credibility (grand average).



The role of the penalty

The penalty encodes assumptions on the coefficients structure

In Akur8 GLM, the structure of the **penalty** encodes meaningful **assumptions/priors on the model**.

These assumptions **differ** on whether a variable is **categorical** or **ordinal** [1].

Categorical variable | **Ordinal variable**

The coefficients:

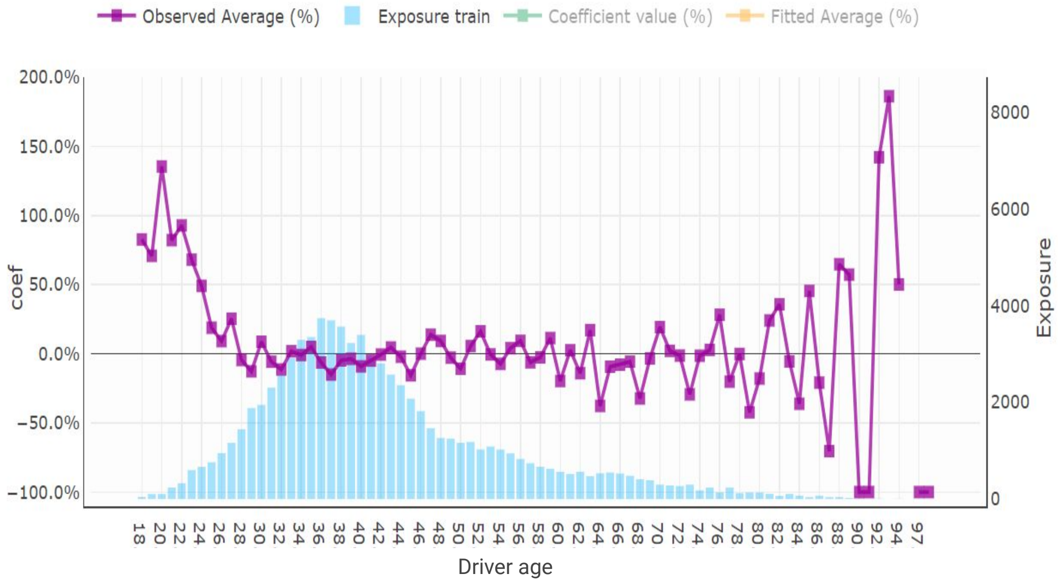
- **Have the tendency to be grouped,** i.e. factors which are close are set to be equal if not significantly different;
- **Require strong evidence to justify big jumps,** important shifts in shape require credible data to be justified.
- Ordinal variables may be modeled as **continuous variables in a traditional GLM**.

[1] Ordinal variables are those where there exists a natural ordering of the levels. The design matrix is a 1/0 encoding equal to a categorical variable.

Example: **Smoothness and *Ordinal* variables**

Raw data contains both **signal** and **noise.**

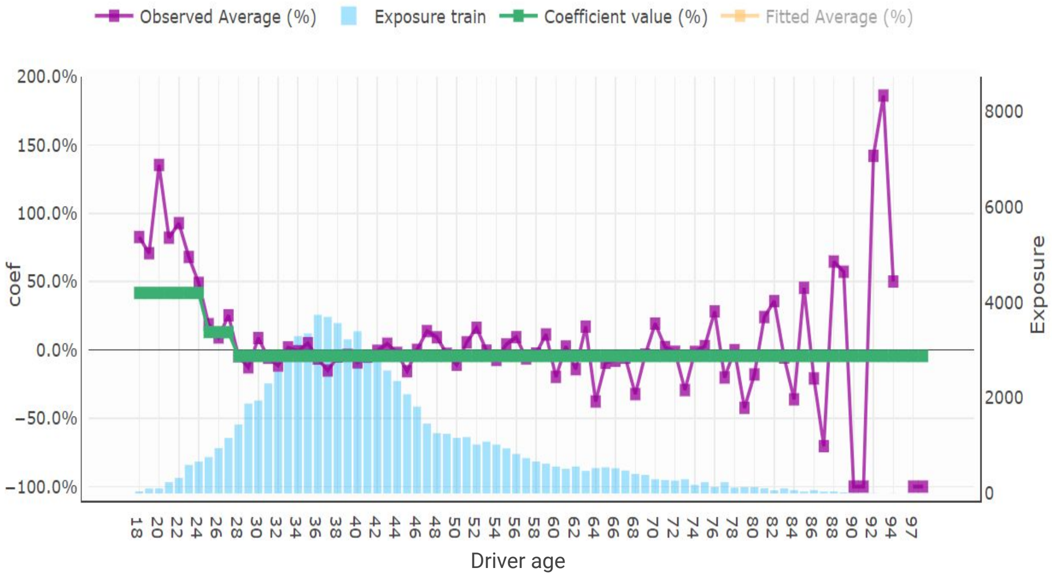
An appropriate model offers a reasonable tradeoff between **robustness** and **sensitivity**.



Example: **Smoothness and *Ordinal* variables**

A **very strong smoothness** leads to models that are **extremely robust**.

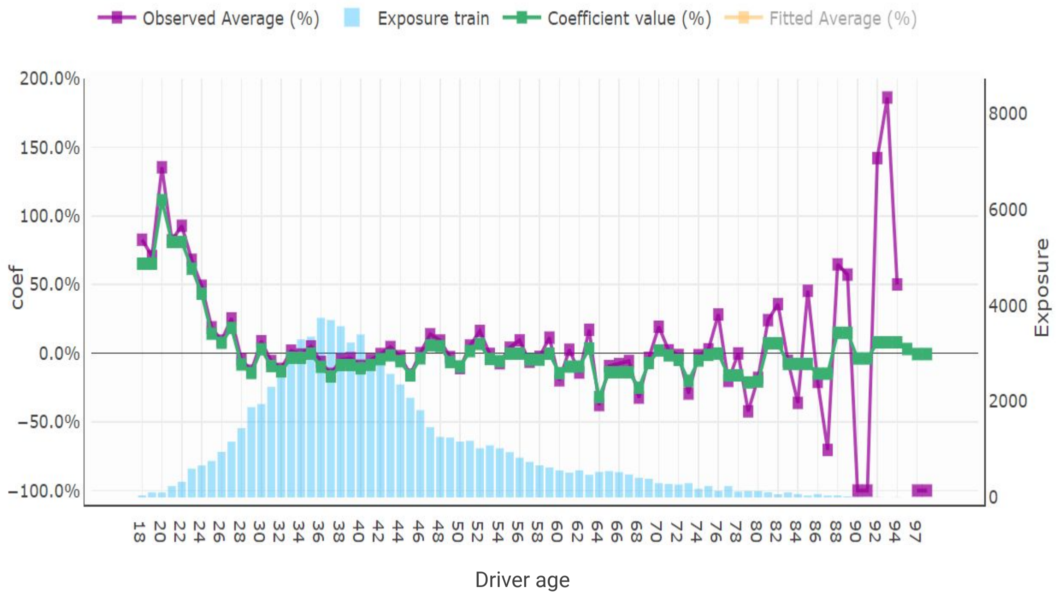
The model segments young drivers versus older drivers, but could segment better between young and older driver age groups.



Example: **Smoothness and *Ordinal* variables**

A very low value of the **smoothness** leads to models that are **extremely sensitive**.

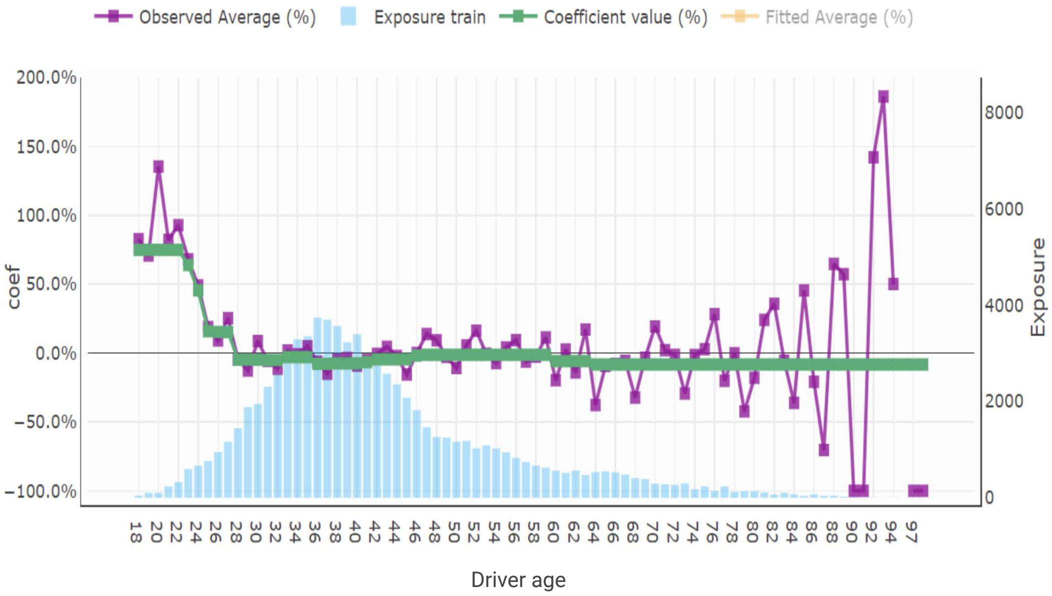
The penalty does not group enough levels, hence the model tries to strictly replicate the observations and the noise associated to it.



Example: **Smoothness and *Ordinal* variables**

A more balanced smoothness leads to **efficient models**.

The model offers a reasonable estimate of the true level of risk for these categories.



Smoothness

What does this parameter do? How should we think about this parameter?

In a typical GLM, the modeler must define “smoothness” by making manual decisions for each variable:

- How many **buckets** for a categorical variable?
- How many **inflection points** in a continuous variable?
- Which **polynomial degree** is appropriate? (squared, cubed, etc.).

In an Akur8 GLM, these are all taken care of by the selection of the Smoothness parameter:

- **Buckets** are automatically created; Higher smoothness → Larger buckets
- **Inflection points** are automatically created; Lower smoothness → More inflection points
- **Nonlinearity** is automatically evaluated. Higher smoothness → Less complex non-linearities

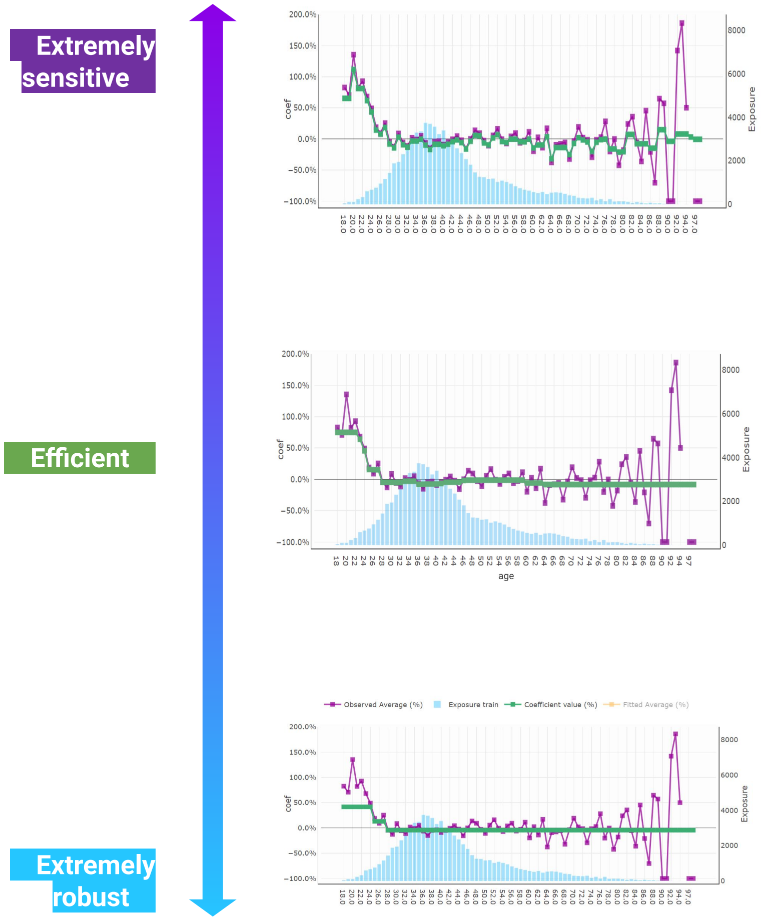
Akur8 allows the modeler to choose the most appropriate smoothness for the model. Evaluating the smoothness in an Akur8 GLM is similar to evaluating the variable transformations made by the modeler in a typical GLM when examining the appropriateness of the model.

Akur8 GLM in practice

The modeling flow

In the Akur8 GLM, the process of feature engineering to detect the non-linearities is greatly reduced. Instead, the modeler should focus on the appropriate choice of the smoothness.

The **Akur8 software builds several models with different values of the smoothness** (7 by default) so that the modeler can investigate the quality of the models and decide whether a lower or higher smoothness is more appropriate [1].



[1] Akur8 provides ability to fine tune this parameter, by **zooming between two** different proposed smoothness **values:** equally spaced values between the two smoothness are tested and the results analyzed.

Methods of evaluation

To choose the smoothness, the modeler **directly evaluates the structure of the model** by looking at the rating factors:

Appropriateness of the model

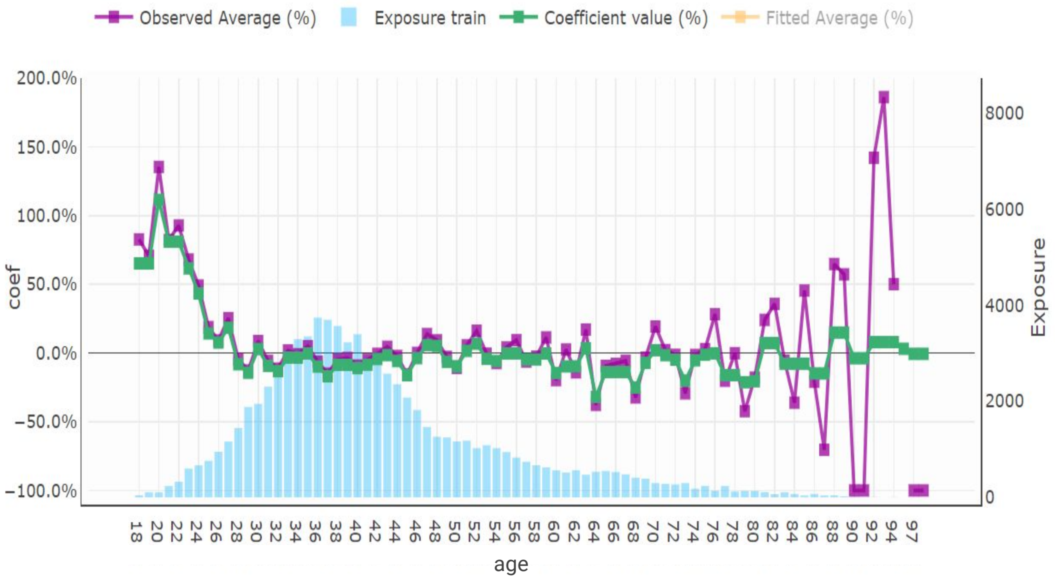
Predictive power of the model

Appropriateness of the model

For this model the smoothness parameter is clearly too low.

The high sensitivity leads to several trend reversal (wiggling) which highlight overfitting.

The modeler should always thoroughly review the trends of the coefficients to make actuarially sound decisions.



Appropriateness of the model

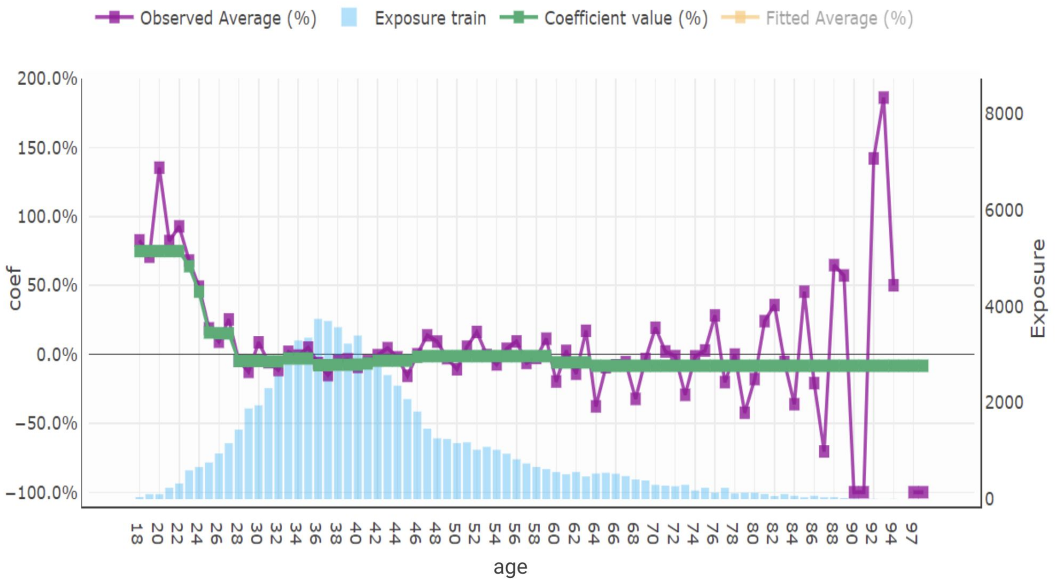
The smoothness is appropriate when the coefficients display an expected and justifiable trend.

This reasonableness must apply to all variables in the model.

If one variable shows a behaviour that is not consistent, the modeler should:

1. Understand whether a problem exists in the underlying data;

2. Increase the smoothness to make the trends more robust.



Predictive power of the model

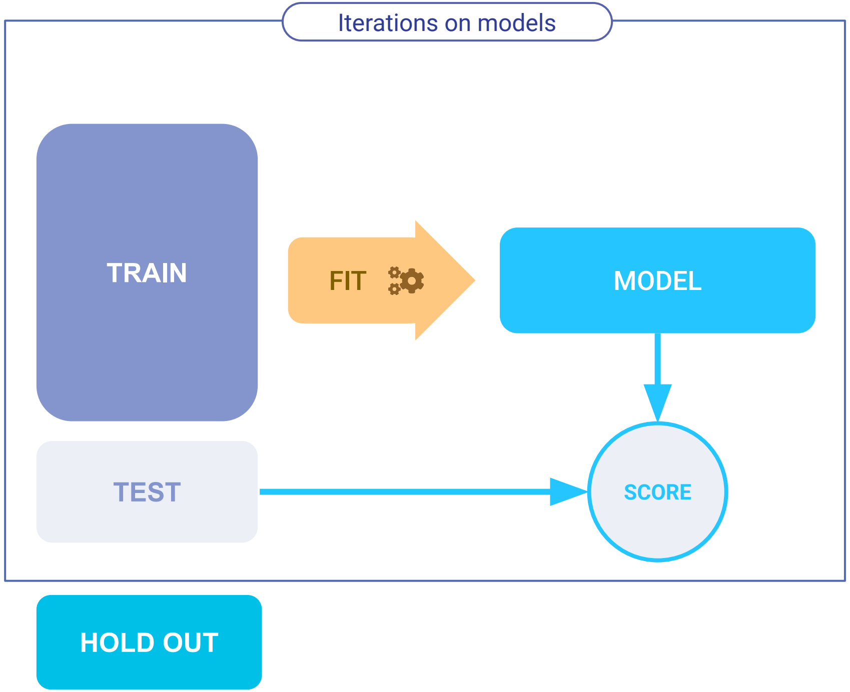
The K-Fold approach

Modeling choices should generalise well to unseen data:

- **Feature Engineering** for GLMs (polynomials, groupings);
- **Smoothness** for Akur8 GLMs.

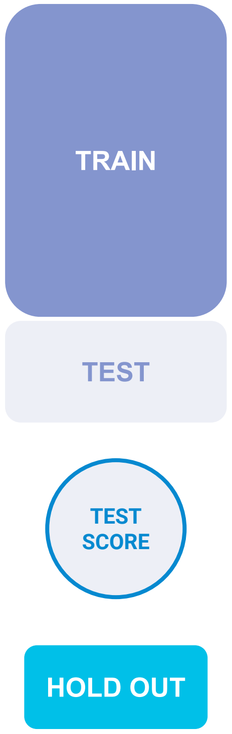
One way to verify the generalisation power, is to **split the data in 3 parts**:

- **Train** to fit the model;
- **Test** to check whether the modeling decisions generalise well;
- **Hold out** is data that will be used only once the modeling decision is final to confirm that the decisions still generalize well.



Generalization Power: Cross Validation

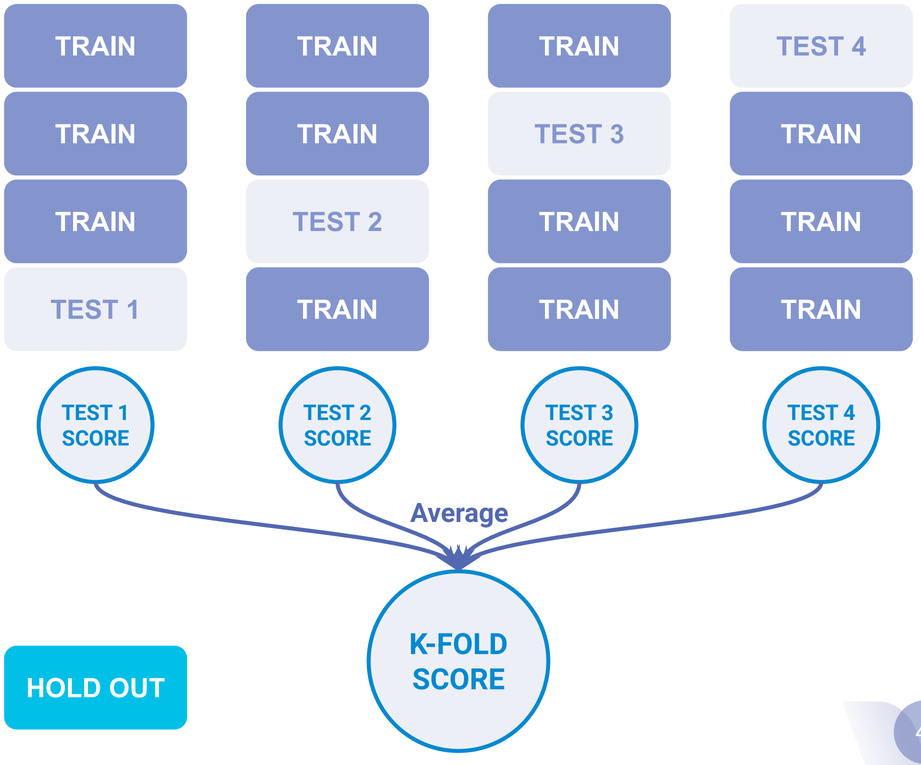
The K-Fold approach



Traditional Train/Test has one test set and relies on a single score.

Cross Validation splits the data to have K test sets and relies on the average of K scores.

Both methods have a holdout dataset



Cross Validation

Summary and impact of the procedure

To summarize, the **cross validation** procedure:

- Robustly **estimates** the predictive power of the model **without using the hold out;**
- Models with **different smoothness values** are compared using cross validation.

The cross validation serves **only to assess the predictive power of the smoothness:**

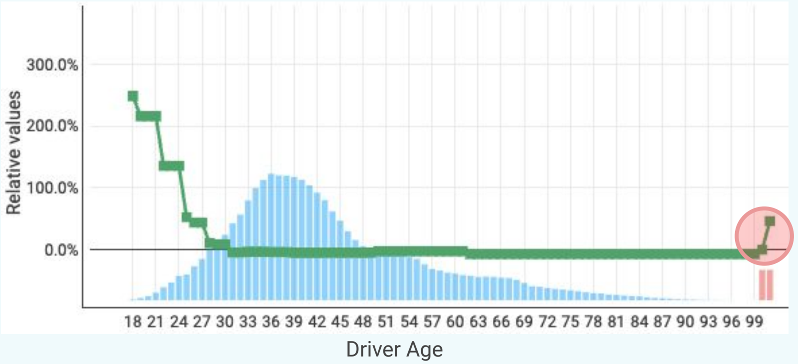
- It is a method of **splitting the data** when **training** the model;
- It **does not** influence the coefficients of the final model since the final model will be fit using all available train data.

Choosing the best tradeoff

What is the right choice of the smoothness?

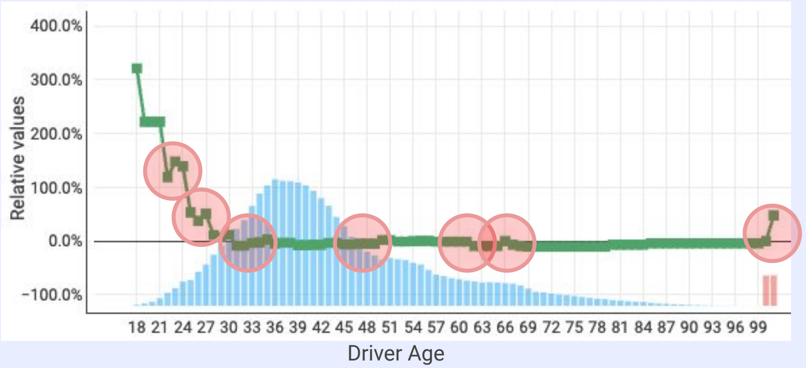
Which model should be selected?

Estimated performance **Gini: 20.5%**
Model is appropriate (1 trend reversal)



Model on the left might lead to **better results** once **deployed in production**.

Estimated performance **Gini: 21%**
Model is less appropriate (> 9 trend reversals)



Model on the right has better predictive power on holdout data, but **may not be actuarially appropriate.**

Regulatory Review

Filing support template

The documentation output of a specific model

The Akur8 software outputs a standard filing support template to verify the actuarial soundness of the model.

The documentation packet follows the best practices as defined by the:

- [Regulatory review of predictive models whitepaper]
- [Speed to Market Book club presentation (April 22)]

The filing template provides best practices for the technical support of the **indicated models.**

The filing template is support for the model for the insurer that may be used in a regulatory filing along with any other documents and forms that are required to be submitted with a filing.

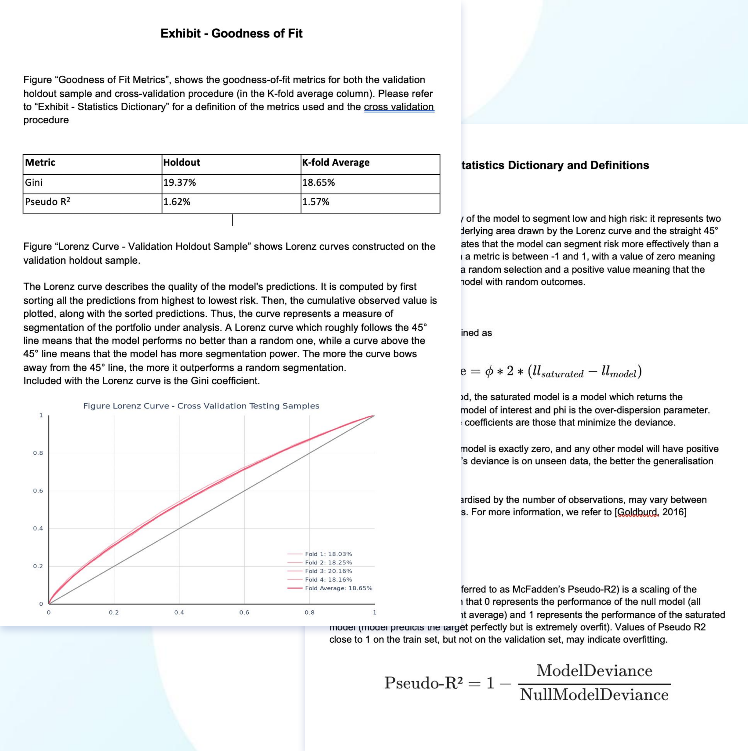


Exhibit Template

The structure of the Exhibits

The Akur8 software provides a list of standardized Exhibits. Exhibits are Word document templates containing narratives, graphs and excel tables.

These exhibits can be roughly grouped into three key topics:

Data

- What and which data was used to develop the models;
- Data Dictionary, Data Summary.

Modeling

- The assumptions made for the models;
- Mathematical Formulas…

Validation

- Evidence that the model is actuarially sound and fits well on holdout data;
- Goodness of fit, Variables Plots,…

Example: **Goodness of Fit Exhibit**

Metrics are displayed in the hold-out and cross validation

The Goodness of Fit Exhibit aims at documenting the ability of the model to generalize on unseen data and not overfit.

It contains metrics and performance plots computed on:

The hold out sample

- The hold out sample is real data and it is not used in the model building process of the software.
- This sample allows model builders and reviewers to confirm that the model generalizes to unseen data.

The cross validation score

- The cross validation score is a metric that guides the modeler when selecting the smoothness parameter.
- This score informs decisions made by the modeler throughout the modeling process.

Example: **Goodness of Fit Exhibit**

Extract of the filing support template

Exhibit - Goodness of Fit

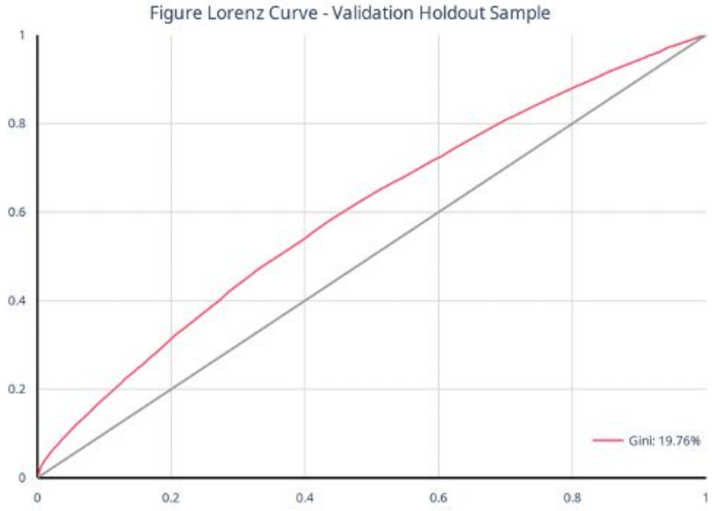
Figure "Goodness of Fit Metrics", shows the goodness-of-fit metrics for both the validation holdout sample and cross-validation procedure (in the K-fold average column). Please refer to "Exhibit - Statistics Dictionary" for a definition of the metrics used and the cross validation procedure

Metric	Holdout	K-fold Average
Gini	19.37%	18.65%
Pseudo R^2	1.62%	1.57%

Figure "Lorenz Curve - Validation Holdout Sample" shows Lorenz curves constructed on the validation holdout sample.

The Lorenz curve describes the quality of the model's predictions. It is computed by first sorting all the predictions from highest to lowest risk. Then, the cumulative observed value is plotted, along with the sorted predictions. Thus, the curve represents a measure of segmentation of the portfolio under analysis. A Lorenz curve which roughly follows the 45° line means that the model performs no better than a random one, while a curve above the 45° line means that the model has more segmentation power. The more the curve bows away from the 45° line, the more it outperforms a random segmentation.
Included with the Lorenz curve is the Gini coefficient.

Figure "Lorenz Curve - Cross Validation Testing Samples" Lorenz curves constructed on the cross-validation testing samples.



Example: **Goodness of Fit Exhibit**

The default metrics used to evaluate the model

Goodness of fit metrics are **computed** on the **Hold Out or** via **Cross Validation**.
The two main metrics to evaluate the Goodness of Fit are:

Gini

- Measures the segmentation power of the model.

Pseudo R^2

- Equivalent to Deviance or Log-Likelihood;
- Measures the distance between observations and predictions;
- It is a "normalised" version of the Deviance:
    - It is equal to 0% when the model is the average;
    - It is equal to 100% when the model is equal to the observations (saturated model);
    - Model with lower deviance will have lower Pseudo R^2.

$$Pseudo\text{-}R^2 = 1 - \frac{ModelDeviance}{NullModelDeviance}$$

Example: **Variable plots**

Evaluating the quality of the trends

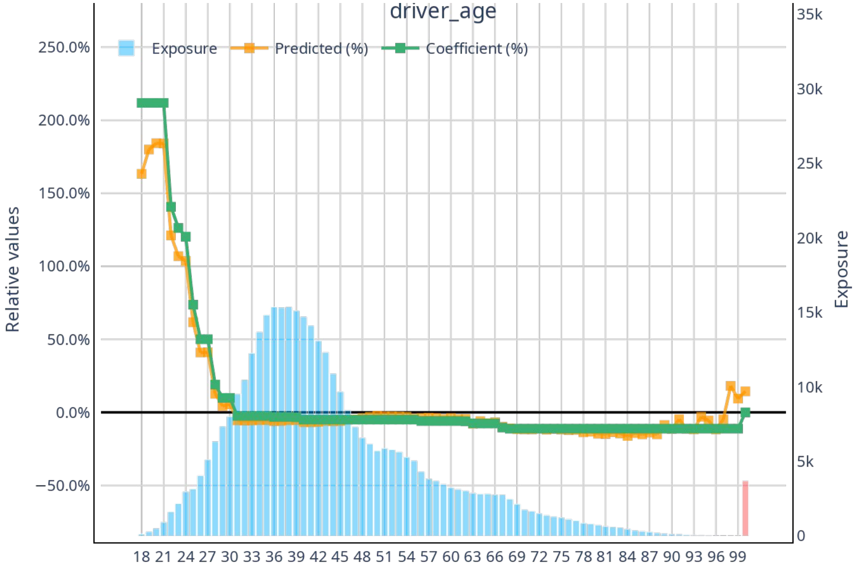
The choice of the smoothness is strongly driven by actuarial judgement around the shape of the rating factors.
Akur8 exports two visualisations for each variable included in the model:

Predicted - Coefficient

- The **green line** represents the **coefficients** / rating factors
- The **orange line** represents the **average prediction** by level, normalized by the grand average.

The plot allows a modeler or reviewer to:

- Verify that the shape of the coefficient is appropriate;
- Identify reversals of the variables with the predictions of the model.



Example: **Variable plots**
Evaluating the quality of the trends

The choice of the smoothness is strongly driven by actuarial judgement around the shape of the rating factors.
Akur8 exports two visualisations for each variable included in the model:

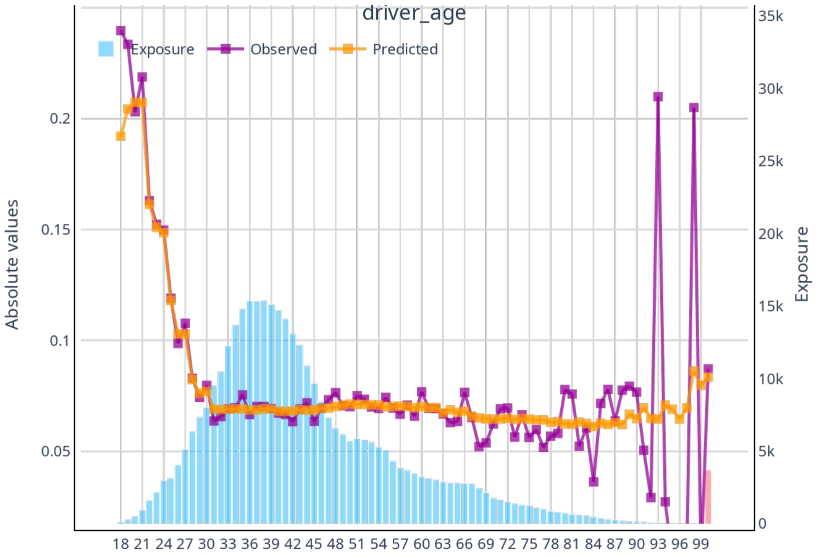
Predicted - Observed

- The **purple line** represents the average observed by level.
- The **orange line** represents the average prediction by level.

The plot allows a modeler or reviewer to:

- Prove the correlation between predictions and losses.

Note that the **predicted** is not the same as the **rating factor** on the prior slide since the prediction takes into account other rating variables.



Example: **Variable significance**

Assessing the soundness of the selections

Variables which are not predictive of losses should not be included in indicated models.

To prove the soundness of the selection, significance tests may be required on the filing support.

Akur8 GLMs template offers by default two indicator of significance at variable level:

- Spread 100/0;
- Spread 95/5.

Example: **Spread 100/0**

A simple yet effective to rank variables by importance

Variable spreads provide a simple yet effective way to assess the impact of a variable in a model.

It is computed as:

$$\text{Spread}_{100/0} = \frac{\text{Max}_{\text{Coefficient}}}{\text{Min}_{\text{Coefficient}}}$$

The spread is an appropriate estimator of significance for Akur8 GLM as each level has its own proper rating factor.

$$\text{Spread}_{100/0} = 74.48\% = \frac{1.1039}{0.6316} - 1$$

Example: **Spread 95/5**

A robust way to assess importance of the variables

The Spread 100/0 is sensitive to outliers.

The Spread 95/5 provides robust estimations of the variable importance.

1. The modalities are sorted by the value of their coefficient;
2. Coefficients below 5% and above 95% exposure are ignored;
3. The spread is computed on the remaining coefficients.

$$\text{Spread}_{95/5} = 34.03\% = \frac{1.1039}{0.8236} - 1$$

Spread and model review

Spreads assist reviewer on assessing model’s soundness

Analysis of the spread support the reviewer for:

- **Detecting outliers** / spot overfitting:
    - When the **difference** between Spread 100/0 and Spread 95/5 is **important**, the **variables should be analysed** to find the levels driving the differences
- Detecting **questionable** variables:
    - Variables whose spread is lower than a selected percentage may require additional narrative to be proved sound. This selection may differ based on the data and line of business;
    - Variables whose spread is unusually high compared to other rating factors may highlight problems with the data.

CONCLUSION

The software

Akur8 provides software for the specific use case of ratemaking

We believe that through collaboration and education, the Akur8 GLM modeling technique can be considered a valid alternative to GLMs whenever credible and non-linear estimates are appropriate modeling assumptions.

Akur8 provides a cloud based **software** to actuaries and ratemaking modelers to build auditable models with efficiency and transparency.

Akur8 GLM methodology

The Akur8 GLM through the choice of the smoothness, develops models which are both sound (by blending credibility) and predictive (by fitting non-linear effects).

Filing support documentation

The filing support template aims at bringing down the barriers that practitioners can have to build support that follows the best practices of the NAIC.

REFERENCES

AKUR8

References

- Fry, Taylor. "A discussion on credibility and penalised regression, with implications for actuarial work." (2015).
- M.Casotto et al. “Credibility and Penalized Regression” (2022)
- Hastie, Trevor, and Junyang Qian. "Glmnet vignette." *Retrieved June* 9.2016 (2014)
- Tibshirani, Robert. "Regression shrinkage and selection via the lasso." (1996)
- Goldburd, Mark, et al. "Generalized linear models for insurance rating." Casualty Actuarial Society, CAS Monographs Series 5 (2016).