LocalGLMnet: A Deep Learning Architecture for Actuaries

Insurance Data Science Conference, Milan

17th June 2022

Dr. Jürg Schelldorfer*, Actuary SAA

Senior Actuarial Data Scientist, Swiss Re

Chair of the «Data Science» working group, Swiss Association of Actuaries (SAA)

*Tutorial co-authored with M.V. Wüthrich; research articles published by M.V. Wüthrich and R. Richmann, see references.

Disclaimer

The opinions expressed in this presentation are those of the author only. They are inspired by the work that the author is doing for both Swiss Re and the SAA, but they do not necessarily reflect any official view of either Swiss Re or the SAA.

Table of Content

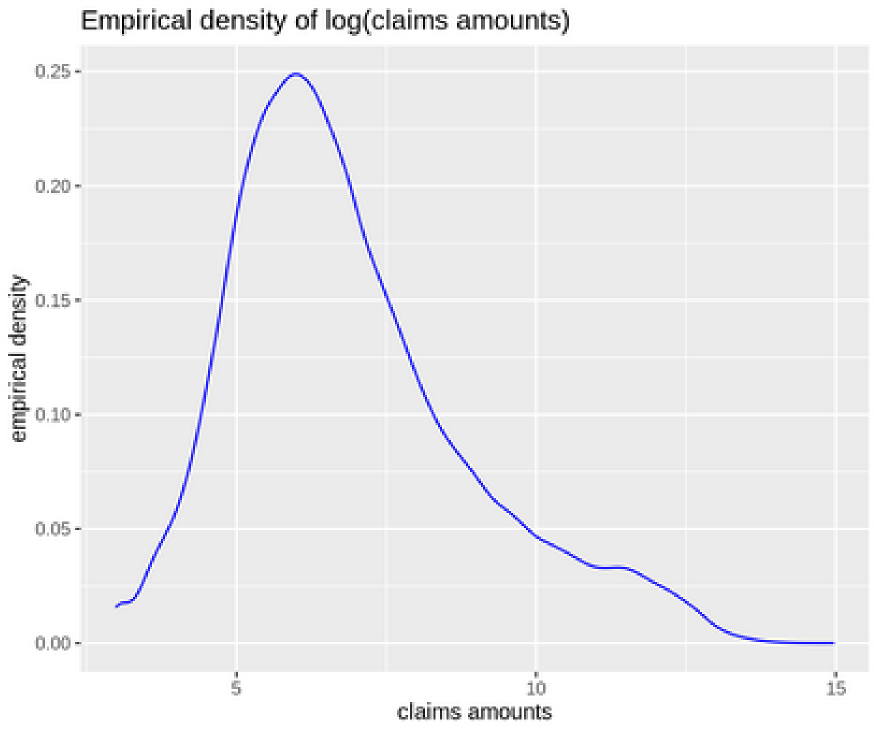
Data

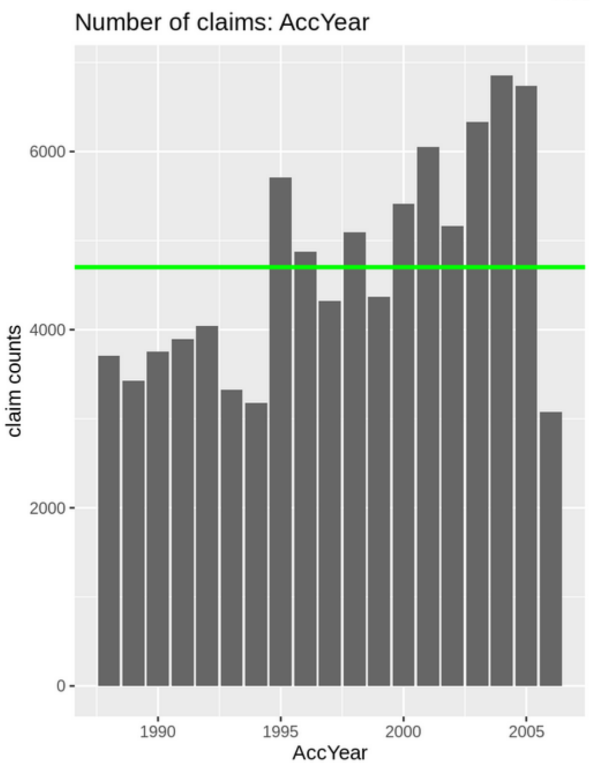
Worker's compensation claims data (1/2)

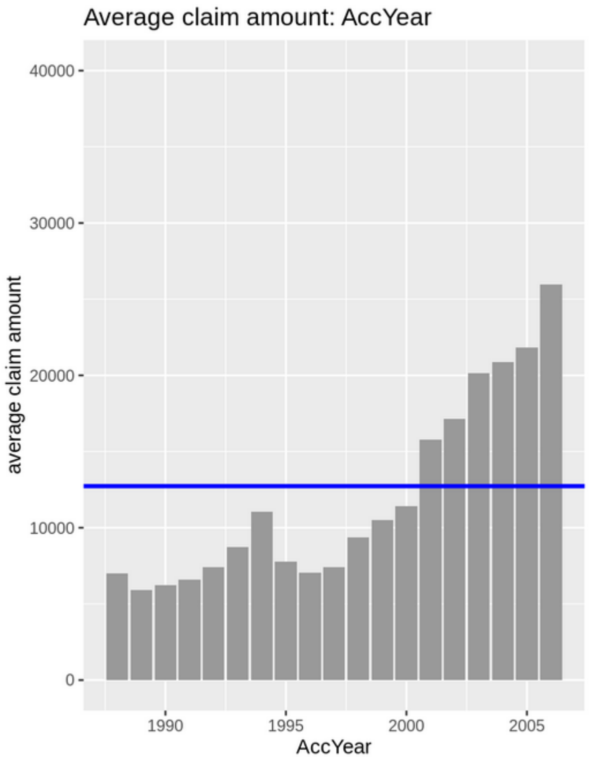
```
## 'data.frame':    89332 obs. of  44 variables:
##  $ ClaimNumber         : chr  "WC9730090" "WC5305468" "WC6197732" "WC8318363" ...
##  $ Age                 : num  53 37 18 34 26 26 28 63 49 29 ...
##  $ Gender              : Factor w/ 2 levels "F","M": 1 1 2 2 2 2 2 2 2 1 ...
##  $ MaritalStatus       : Factor w/ 3 levels "M","S","U": 1 3 2 2 3 3 2 2 1 2 ...
##  $ DependentChildren   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ DependentsOther     : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ WeeklyPay           : num  683 535 250 544 435 ...
##  $ PartTimeFullTime    : Factor w/ 2 levels "F","P": 1 1 1 1 1 1 1 1 1 1 ...
##  $ HoursWorkedPerWeek  : num  40 38 38 38 37 20 40 38 38 60 ...
##  $ DaysWorkedPerWeek   : num  5 5 5 5 5 4 5 5 5 5 ...
##  $ ClaimDescription    : chr  "BUFFING OUT DISC CUTTING RIGHT UPPER SWOLLEN HEAD" "FELL HE
AD ON STEP PULLED TWISTED SOFT TISSUE RIGHT ANKLE" "KNIFE SLIPPED AND METAL LACERATION HEAD"
"SLIPPED DOWN ROCK SOFT TISSUE INJURY RIGHT INDEX KNEE" ...
##  $ InitialCaseEstimate : num  20000 7500 425 70000 25500 ...
##  $ Claim               : int  3139046 2895985 2049604 2023994 1669706 1610744 1452597 1423
497 1339464 1282092 ...
##  $ AccDate             : Date, format: "2001-09-10" "1994-07-18" ...
##  $ AccDay              : int  5005 2394 1023 5762 4991 6189 6318 5565 5766 3529 ...
##  $ AccYear             : int  2001 1994 1990 2003 2001 2004 2005 2003 2003 1997 ...
##  $ AccMonth            : int  9 7 10 10 8 12 4 3 10 8 ...
##  $ AccWeekday          : num  1 1 2 2 1 2 5 1 6 2 ...
##  $ AccTime             : int  12 7 13 12 13 10 17 11 16 12 ...
##  $ RepDate             : Date, format: "2001-10-08" "1994-08-19" ...
##  $ RepDay              : int  5033 2426 1165 5778 5012 6219 6330 5575 5782 3539 ...
##  $ RepYear             : int  2001 1994 1991 2003 2001 2005 2005 2003 2003 1997 ...
```

- This dataset describes realistic, synthetically generated worker compensation insurance claims.
- Along the ultimate financial losses, each claim is described by the initial case estimate, date of accident and reporting date, a text describing the accident and demographic info on the worker.
- The dataset was kindly created and provided by Colin Priest. While similar, it is not identical to the dataset used in www.kaggle.com/c/actuarial-loss-estimation.
- https://www.openml.org/d/42876

Worker's compensation claims data (2/2)







GLM and FNN

Generalized Linear Model (GLM)

g strictly monotone link function, regression function

$$\boldsymbol{x} = \left(x_1, \dots, x_q\right) \mapsto g(\mu^{GLM}(\boldsymbol{x})) = \beta_0 + \sum_{j=1}^{q} \beta_j x_j,$$

where the regression parameter $\boldsymbol{\beta} = (\beta_0, \dots, \beta_q)$ is estimated by MLE.

The concrete structure of $\boldsymbol{x}$ is an input (e.g. age as numerical feature or age bins). Often manual feature engineering.

Feedforward Neural Network (FNN)

Regression function

$$\boldsymbol{x} = \left(x_1, \dots, x_q\right) \mapsto g(\mu^{FNN}(\boldsymbol{x})) = \beta_0 + \sum_{j=1}^{q} \beta_j z_j^{(d:1)}(\boldsymbol{x}),$$

where $\boldsymbol{x} \mapsto \boldsymbol{z}^{(d:1)}(\boldsymbol{x})$ is a network of depth d.

The concrete structure of $\boldsymbol{z}$ is learned by the network from $\boldsymbol{x}$.

Generalized Linear Model (GLM), revisited

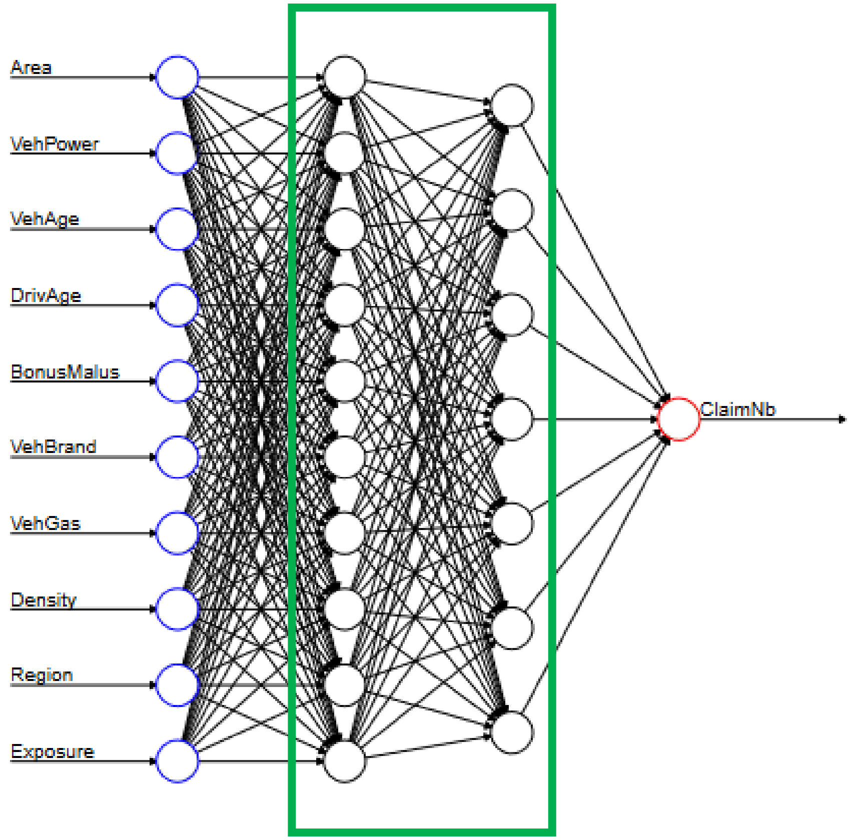
Regression function

$$\boldsymbol{x} = (x_1, \dots, x_q) \mapsto g(\mu^{GLM}(\boldsymbol{x})) = \beta_0 + \sum_{j=1}^{q} \beta_j z_j^{(d:1)}(\boldsymbol{x}),$$

where $\boldsymbol{x} \mapsto \boldsymbol{z}^{(d:1)}(\boldsymbol{x})$ is a network of depth d.

A GLM is the special case of d=1!

GLM vs FNN



d=1: GLM
d=3: FNN

Advantages

- The network learns the representation of the input $\boldsymbol{x}$
- A well-calibrated FNN often outperforms a GLM.

Limitations of FNN

- Not interpretable
- No simple way of variable selection

LocalGLMnet

LocalGLMnet: Definition

FNN of depth d and input and output dimension equal to q.

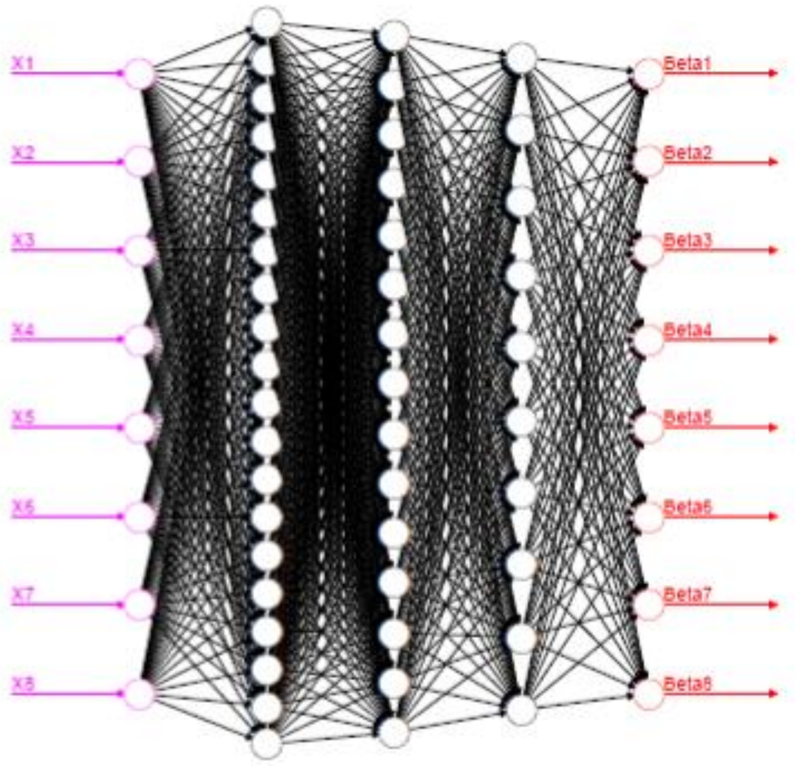
$$\boldsymbol{z}^{(d:1)} : \mathbb{R}^q \to \mathbb{R}^q$$

$$\boldsymbol{x} \mapsto \boldsymbol{\beta}(\boldsymbol{x}) = \boldsymbol{z}^{(d:1)}(\boldsymbol{x}) \quad ,$$

The LocalGLMnet is defined as the decomposition:

$$\boldsymbol{x} \mapsto g\left(\mu^{LocalGLMnet}(\boldsymbol{x})\right) = \beta_0 + \sum_{j=1}^{q} \beta_j(\boldsymbol{x}) x_j$$

LocalGLMnet: Remarks



- LocalGLMnet, because locally around a given $\boldsymbol{x}$, the regression function can be understood as a GLM.

- $\boldsymbol{\beta}(\boldsymbol{x})$ are called regression attentions because they provide more or less attention to specific components of $\boldsymbol{x}$ in the regression function.

LocalGLMnet: Interpretation

Based on:

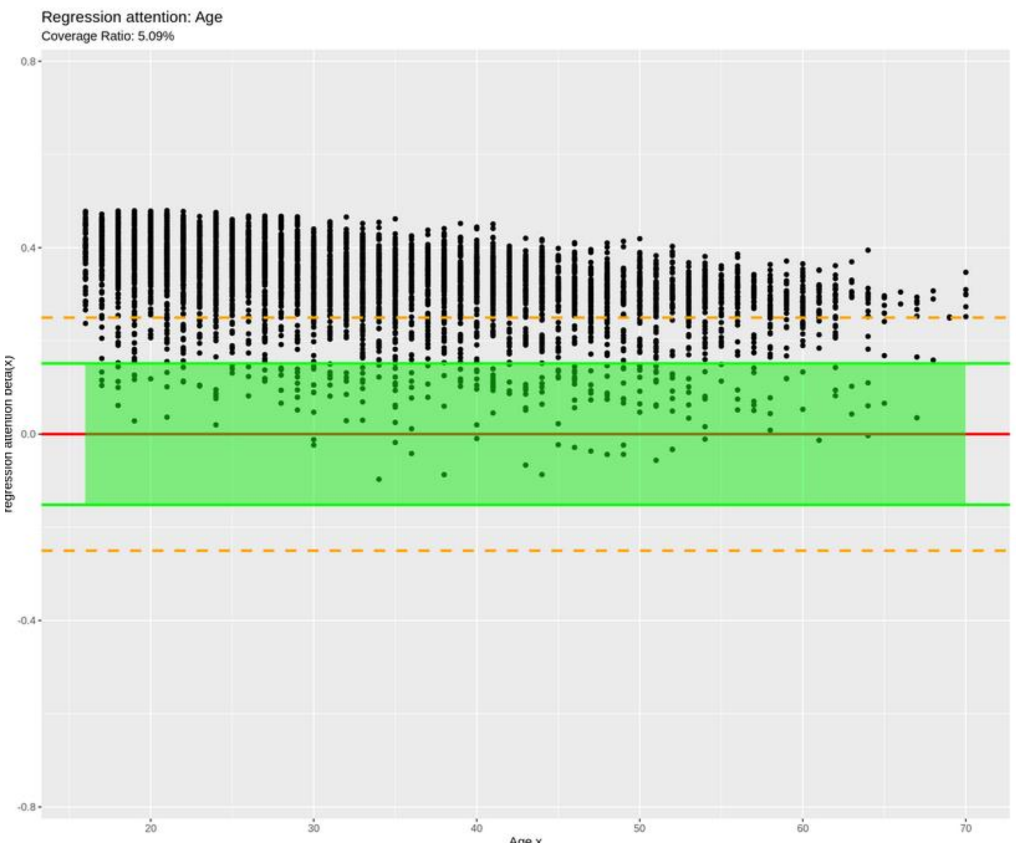
$$\boldsymbol{x} \mapsto g\left(\mu^{LocalGLMnet}(\boldsymbol{x})\right) = \beta_0 + \sum_{j=1}^{q} \beta_j(\boldsymbol{x}) x_j$$

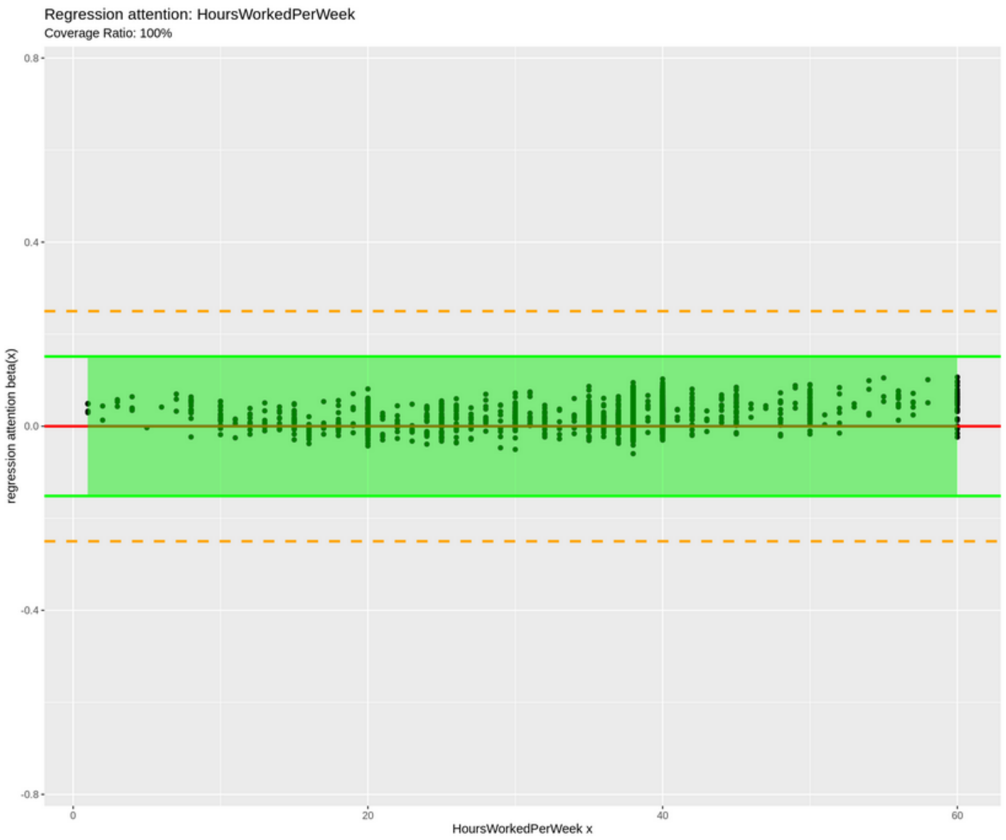
- $\beta_j(\boldsymbol{x}) \equiv 0$: drop term x_j: variable selection
- $VI_j = \frac{1}{n}\sum_{i=1}^{n} \left|\beta_j(x_i)\right|$: feature importance measure
- $\beta_j(\boldsymbol{x}) \equiv \beta_j\left(x_j\right)$: interactions; no interaction with other features $x_{k \neq j}$
- $\beta_j(\boldsymbol{x}) \equiv \beta_j \neq 0$: no covariate dependence, GLM term in x_j

Application

LocalGLMnet: Variable selection

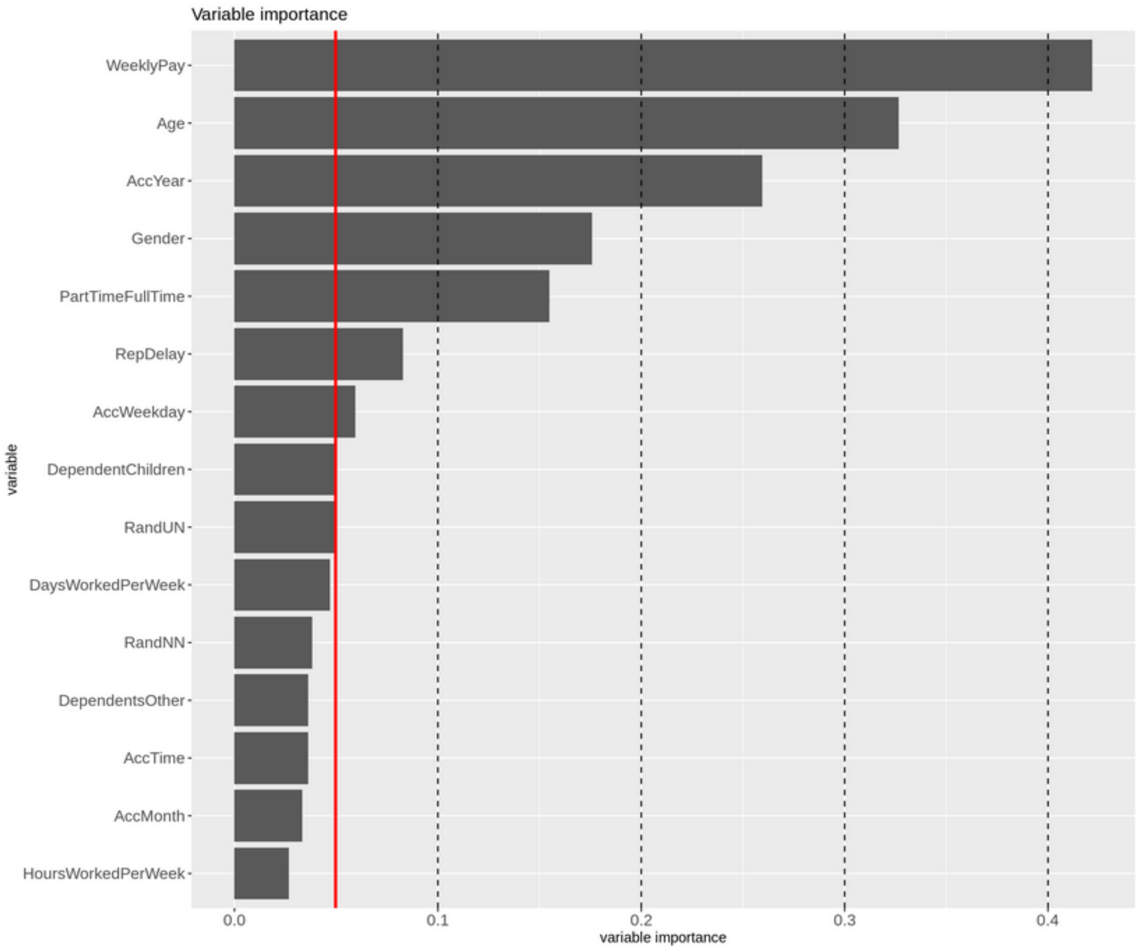
$\beta_j(\boldsymbol{x})$





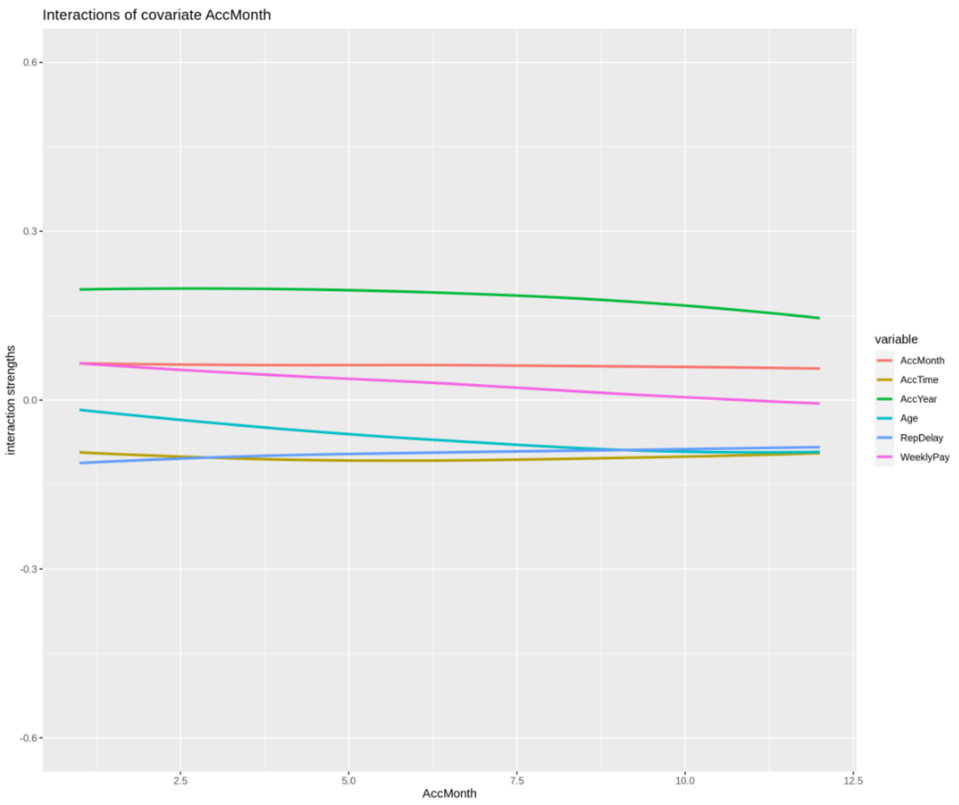
x_j x_j

LocalGLMnet: Feature importance

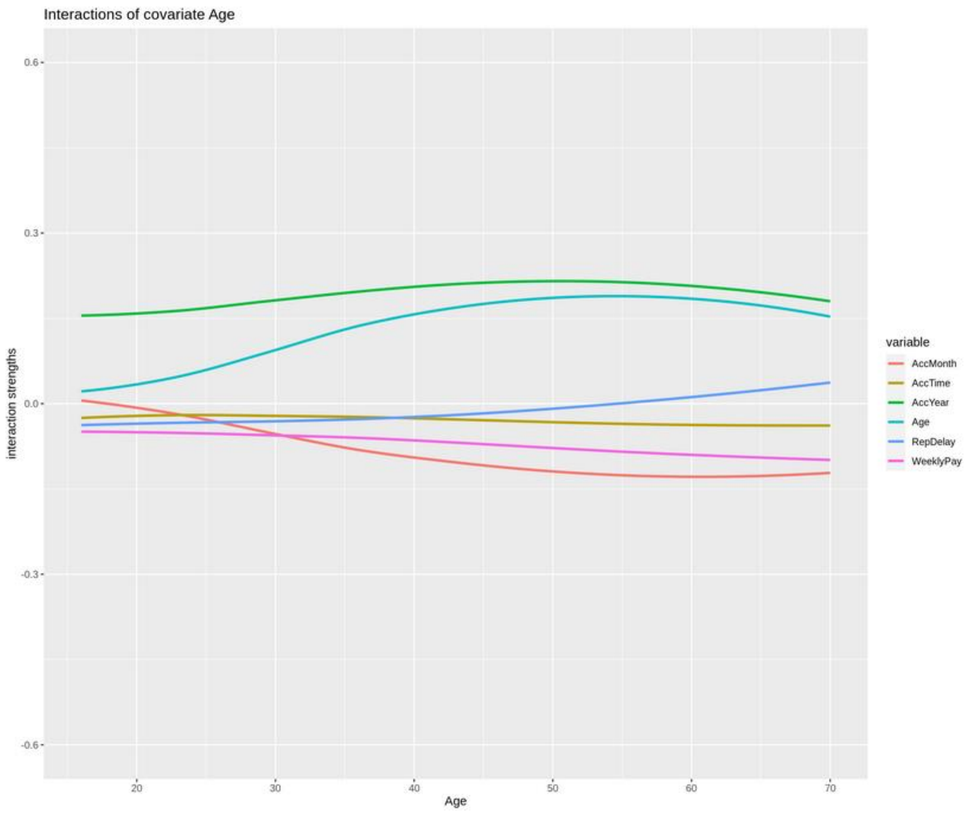


$$VI_j = \frac{1}{n}\sum_{i=1}^{n}\left|\beta_j(x_i)\right|$$

LocalGLMnet: Interactions and linear terms



No interaction of AccMonth with other continuous features



slight interaction between Age and AccYear

Summary

Conclusions

- Worker's Compensation data: https://www.openml.org/d/42876
- GLMs are a special case of FNN, or FNNs are an extension of GLMs
- LocalGLMnet
  - is explainable like a GLM,
  - allows variable selection,
  - allows for feature importance measure,
  - allows to detect interactions.
- Tutorial available here, and corresponding R Notebook here.
- www.actuarialdatascience.org

References

- Richman, Wüthrich (2021). LocalGLMnet: interpretable deep learning for tabular data. SSRN Manuscript, ID 3892015.
- Richman, Wüthrich (2021). LASSO regularization within the LocalGLMnet architecture. SSRN Manuscript, ID 3927187.
- Schelldorfer, Wüthrich (2021). LocalGLMnet: A Deep Learning Architecture for Actuaries. SSRN Manuscrip, ID 3900350.
- M.V. Wüthrich (2022), LocalGLMnet: An Interpretable Deep Learning Architecture. EAA e-Conference on Data Science & Data Ethics, 12th May 2022.

Acknowledgements

People:

- All members of the SAA working group
- Dr. Andrea Ferrario
- Dr. Tobias Fissler
- Dr. Roger Hämmerli
- Mara Nägelin
- Dr. Alexander Noll
- Dr. Simon Renzmann
- Ron Richman
- Dr. Robert Salzmann

Insititutions:

- Swiss Association of Actuaries (SAA)
- RiskLab at ETH Zurich
- MobiLab for Analytics at ETH Zurich

Companies:

- Swiss Re