KU LEUVEN

Neural Network-based models for Non-Life Insurance Pricing

Master Thesis Presentation

Freek Holvoet

28 May 2021

Promotor: Prof. Dr. Katrien Antonio
Work leader: ir. Roel Henckaerts

Introduction

Pure premium calculation

- Technical premiums defined by the expected loss
- Popular frequency-severity approach in P&C pricing:

$$\pi = \mathbb{E}\left(\frac{L}{e}\right) \overset{\text{indep.}}{=} \mathbb{E}\left(\frac{N}{e}\right) \times \mathbb{E}\left(\frac{L}{N} \,\middle|\, N > 0\right) = \mathbb{E}(F) \times \mathbb{E}(S)$$

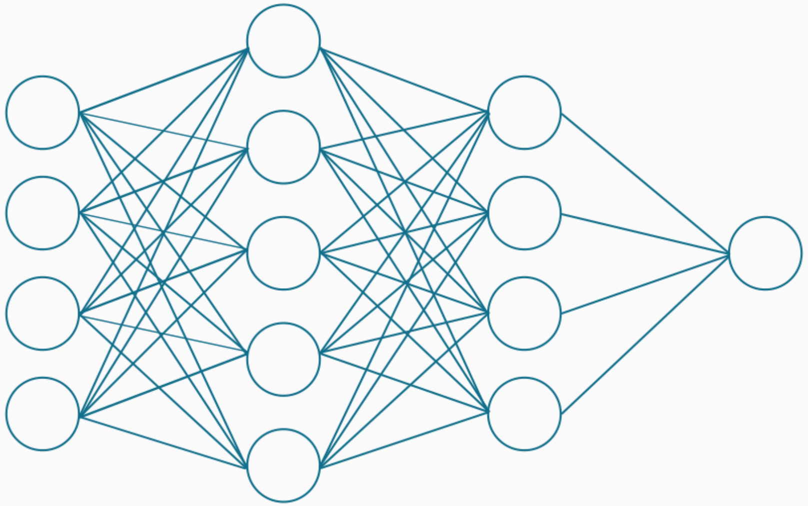
Predictive models

Many options exist, for example: GLM, GAM, decision trees, ensemble methods, Tweedie compound Poisson-gamma models, neural networks, support vector machines.

- Binned GLM: Henckaerts et al. (2018) uses GAMs together with tree-based clustering, to bin continuous variables for use in a GLM
- Gradient Boosting: Henckaerts et al. (2020) uses GBMs for both frequency and severity modelling

Neural Networks

Fully-connected feed-forward neural networks

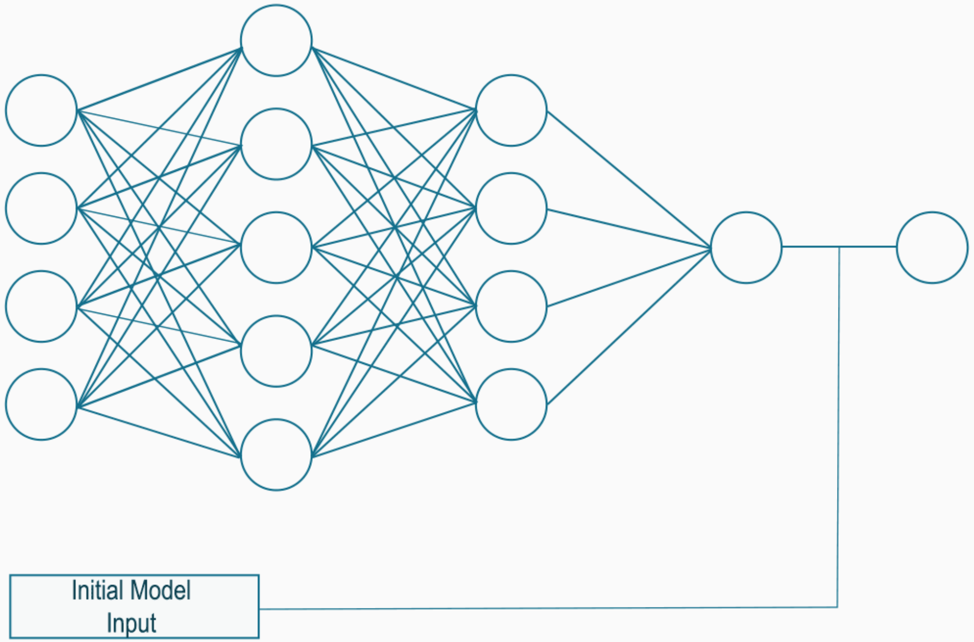


Each layer outputs a weighted combination of the previous layer, with an activation function $\boldsymbol{\sigma}$:

$$\mathbf{layer}_{\boldsymbol{i}} = \boldsymbol{\sigma}(\mathbf{W}\,\mathbf{layer}_{\boldsymbol{i-1}} + \boldsymbol{\beta})$$

Combined Actuarial Neural Networks

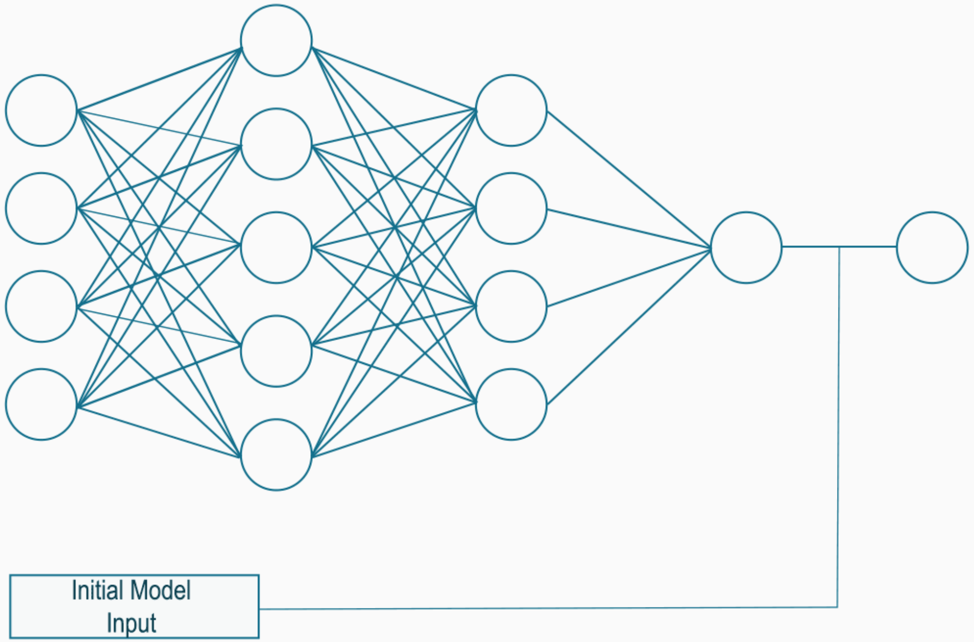
Combined Actuarial Neural Networks (**CANN**), proposed by Scheldorfer and Wutrich (2019)



An **initial model** is added with a skip connection to the output layer. The neural network part is called the **adjustment model**.

Combined Actuarial Neural Networks

Combined Actuarial Neural Networks (**CANN**), proposed by Scheldorfer and Wutrich (2019)



Fixed CANN setup

$$f^{fixed}\left(\boldsymbol{x}_i, \hat{y}_i^{(in)}\right) = \exp\left(\ln\left(\hat{y}_i^{(in)}\right) + \hat{y}_i^{(adj)}\right)$$

Flexible CANN setup

$$f^{flexible}\left(\boldsymbol{x}_i, \hat{y}_i^{(in)}\right) = \exp\left(\begin{bmatrix} w_1 & w_2 \end{bmatrix} \cdot \begin{bmatrix} \ln(\hat{y}_i^{(in)}) & \hat{y}_i^{(adj)} \end{bmatrix} + \beta\right)$$

Bias regularization

Balance property

In a model $\boldsymbol{f}(\boldsymbol{x})$ with perfect balance:

$$\sum_i f(\boldsymbol{x}_i) = \sum_i y_i.$$

Important in insurance: premiums are defined on portfolio level

GLM bias regularization

Wutrich (2020) proposes the use of GLM bias regularization to restore balance in neural network models.

- GLM with canonical link possesses the balance property
- Using the last hidden layer of a neural network as input in a GLM to restore balance in the neural network

Model setup

Different models are compared

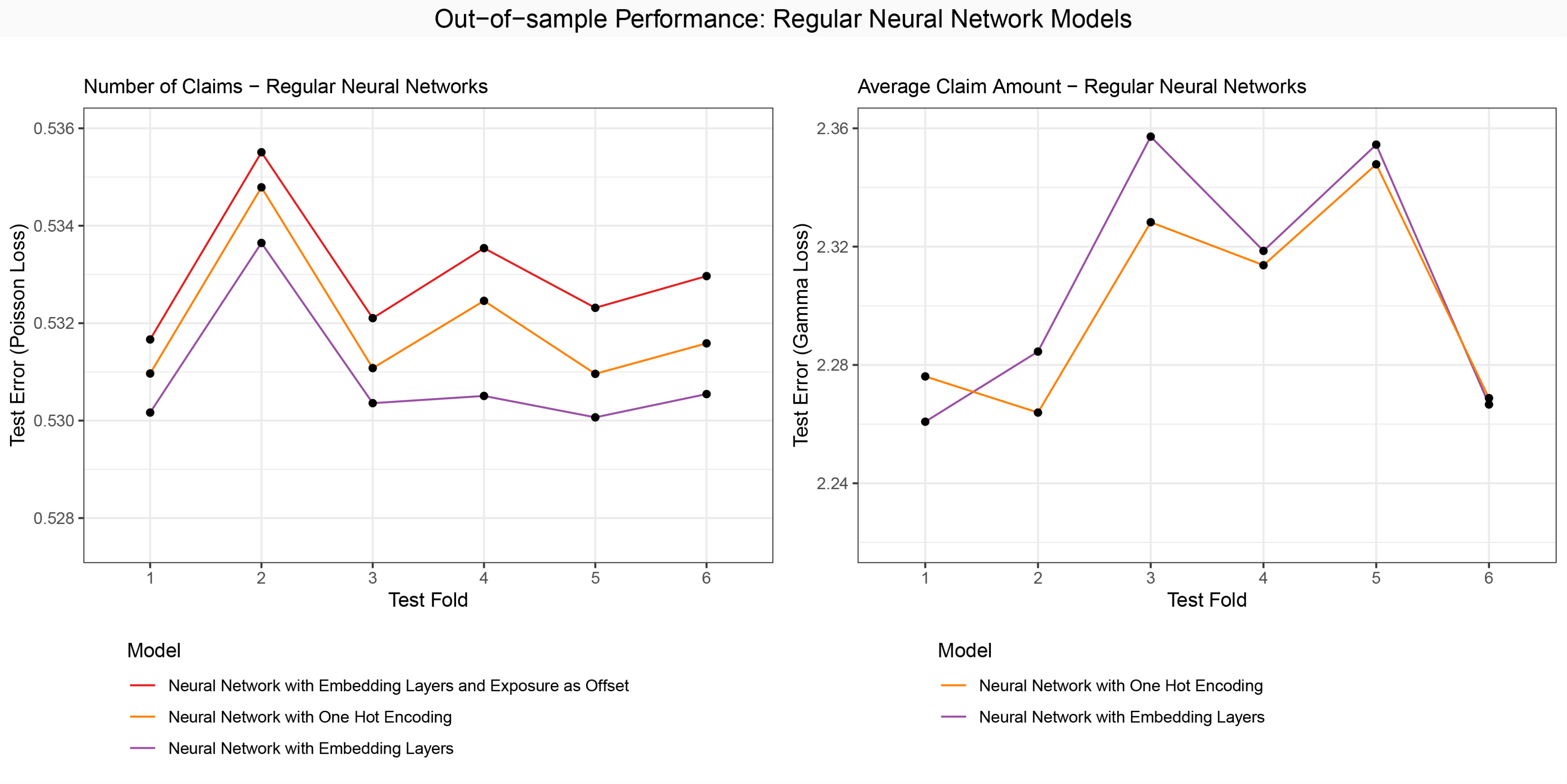
- Regular neural networks
    - One-hot encoding
    - Embedding layers
    - Exposure as offset versus input variable
- CANN models
    - Binned GLM input, fixed and flexible
    - GBM input, fixed and flexible
- All models with and without GLM bias regularization

Training the models

- Loss functions
    - Poisson deviance for frequency model
    - Gamma deviance for severity model
- Tuning strategy
    - K-times $K-1$-fold cross-validation
    - Sequential tuning
- K out-of-sample performances with random early stopping set

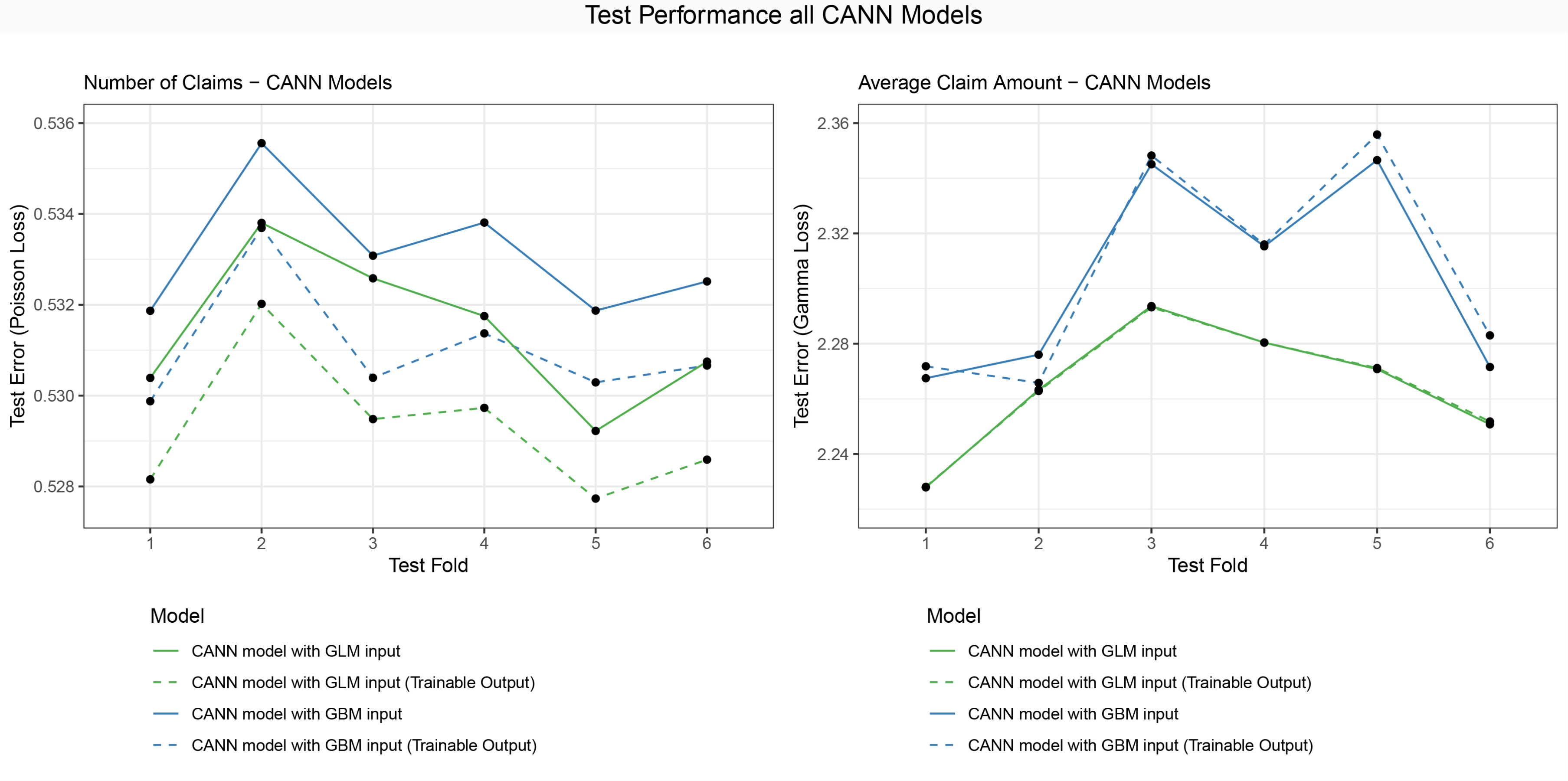
Out-of-sample performance

Neural network results



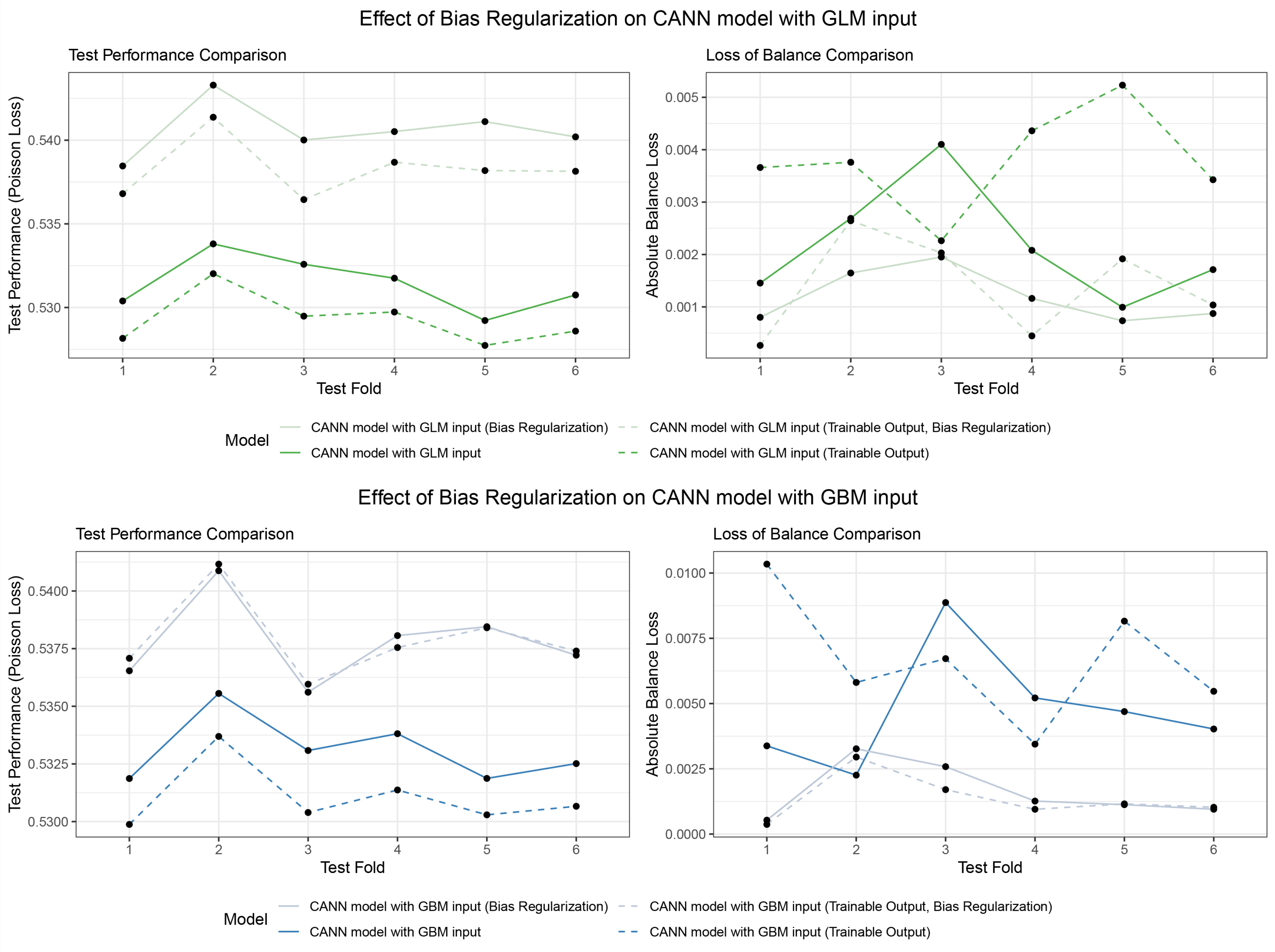
Out-of-sample performance

CANN model results



Out-of-sample performance

Effect of bias regularization



Model interpretational tools

Variable importance plot

- Permutate each variable ℓ
- Calculate the difference in prediction between the unpermutated and the permutated sample data set

$$\mathrm{VIP}_\ell = \frac{1}{n}\sum_{i=1}^{n}\left(f_{\text{model}}(\boldsymbol{x}_i) - f_{\text{model}}\left(\boldsymbol{x}_i^{\text{perm},\ell}\right)\right)$$

Partial dependence plot

- Fix the value of variable ℓ
- Look at the average prediction
- Calculate the average predictions over the full range of possible values

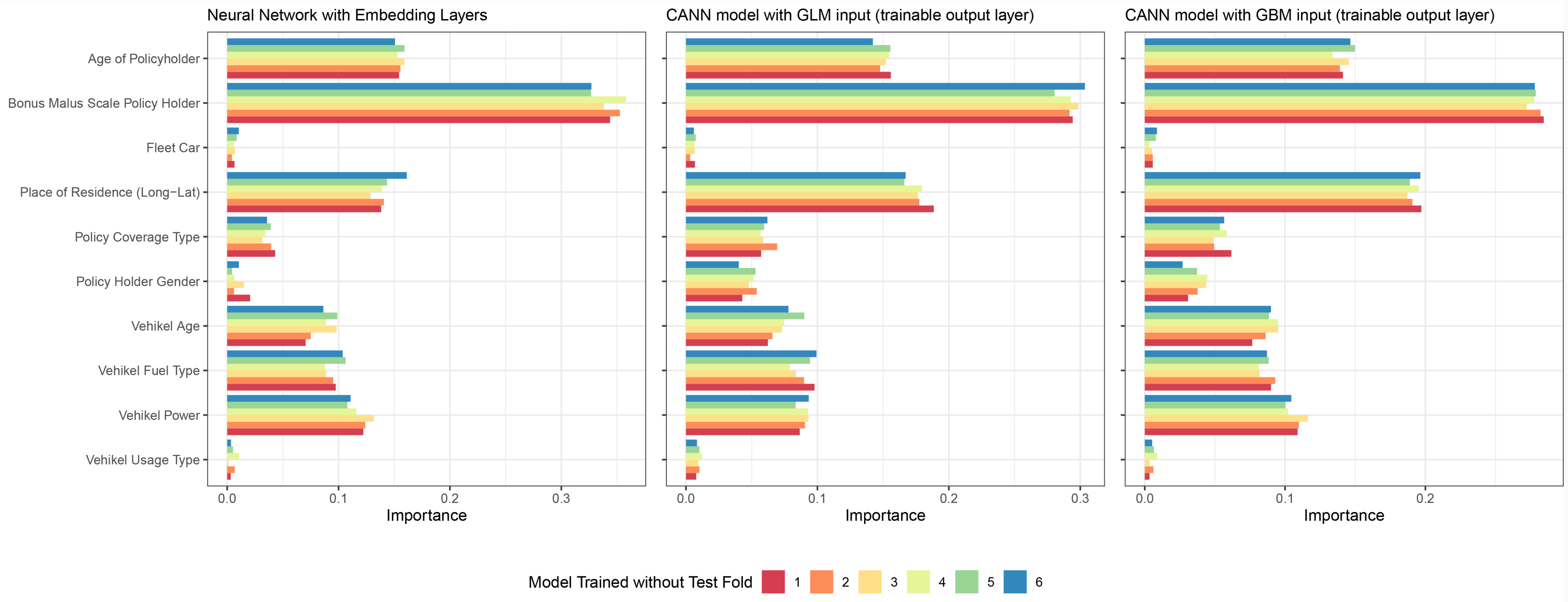
$$\mathrm{PDP}(x_\ell) = \frac{1}{n}\sum_{i=1}^{n} f_{\text{model}}(\boldsymbol{x}_\ell, \boldsymbol{x}_i^*)$$

Reference: Molnar (2019)

Model interpretational tools

Variable importance plot

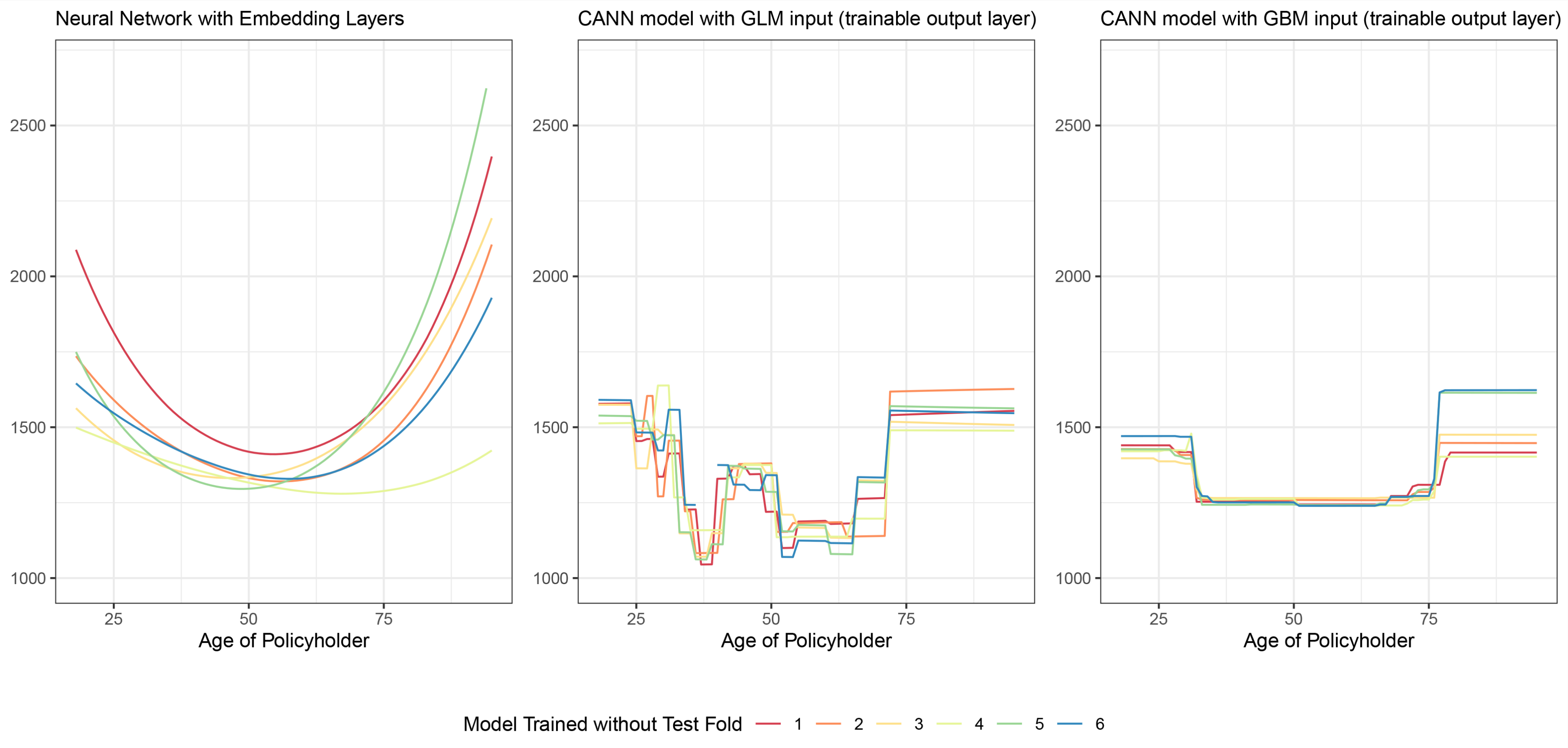
A permutational approach to asses the importance of each variable in the claim frequency model.



Model interpretational tools

Partial dependence plot: Policyholder age

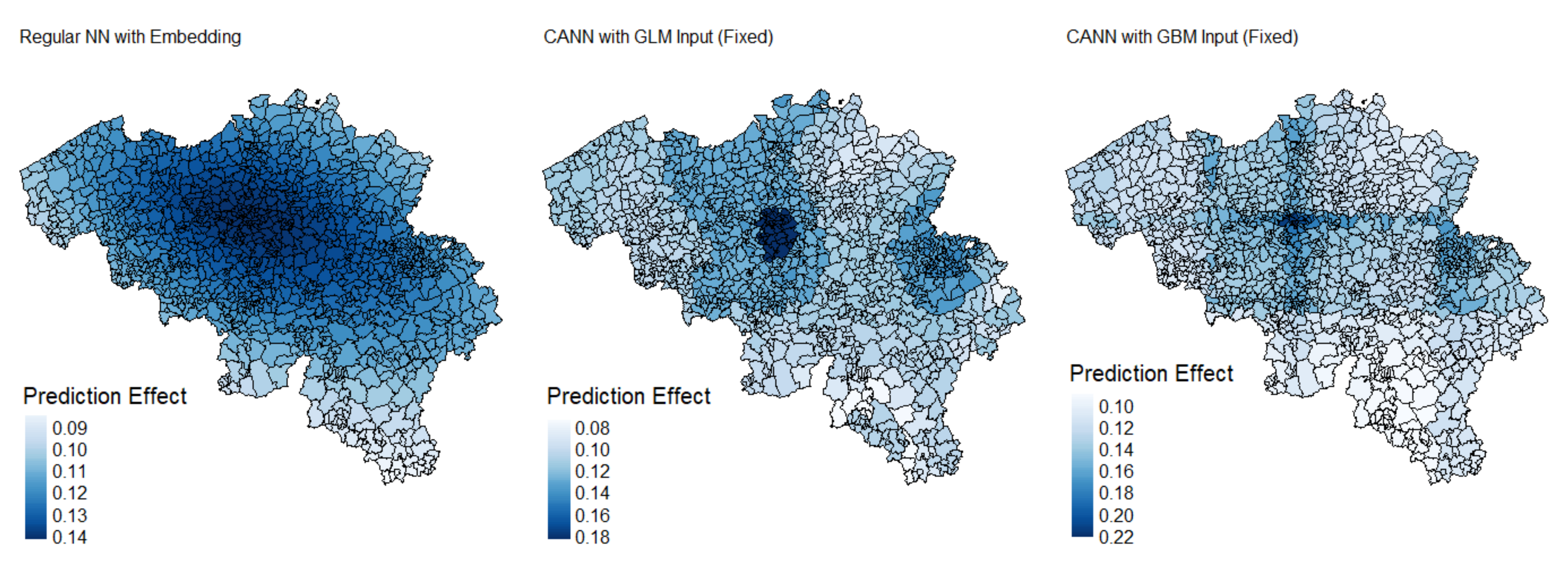
Effect of the policyholder age on the predicted claim severity.



Model interpretational tools

Partial dependence plot: Postal code

Effect of the postal code on the predicted number of claims.



References

Roel Henckaerts, Katrien Antonio, Maxime Clijsters, and Roel Verbelen. A data driven binning strategy for the construction of insurance tariff classes. *Scandinavian Actuarial Journal*, 2018(8):681–705, 2018.

Roel Henckaerts, Marie Pier Cote, Katrien Antonio, and Roel Verbelen. Boosting Insights in Insurance Tariff Plans with Tree-Based Machine Learning Methods. North American Actuarial Journal, pages 1–31, 2020.

Christoph Molnar. *Interpretable Machine Learning*. christophm.github.io/interpretable-ml-book, 2019.

Jurg Schelldorfer and Mario V. Wuthrich. Nesting Classical Actuarial Models into Neural Networks. *SSRN Electronic Journal*, pages 1–27, 2019.

Mario V. Wuthrich. Bias regularization in neural network models for general insurance pricing. *European Actuarial Journal*, 10(1):179–202, 2020.

Slides created with the R package xaringan.
Presentation template from https://github.com/katrienantonio/hands-on-machine-learning-R-module-1