Institute and Faculty of Actuaries

Case Reserving in Non-Life Practice using Individual Data and Machine Learning

Marco Aleandri, PhD, FRM

Univ. “La Sapienza”, Rome

GIRO 2018, Birmingham

21 September 2018

Some Analytics in Actuarial Practice

- Renewal Rate Prediction in Third-Party Liability Insurance
- Case Reserving in Automobile Accident Insurance
- Dynamic Policyholder Behaviour in Life Insurance

Institute and Faculty of Actuaries

Some Analytics in Actuarial Practice

- Renewal Rate Prediction in Third-Party Liability Insurance
- Case Reserving in Automobile Accident Insurance
- Dynamic Policyholder Behaviour in Life Insurance

Institute and Faculty of Actuaries

Agenda

- Traditional Methods vs. Machine Learning
- Australian Bodily Injury Data: a case study
- Closing Delay Estimation
- Claim Amount Prediction
- Reserve Evaluation
- References
- Q&A

Institute and Faculty of Actuaries

Institute and Faculty of Actuaries

Traditional Methods vs. Machine Learning

21 September 2018

What we *learn*…

- Run-off triangles (Chain-Ladder, Fisher-Lange, Mack, etc.)
  - Easy to adapt to different branches
  - Easy to enhance for different purposes
  - Easy to visualize
  - No computational cost

- Regression methods (Gamma, Pareto, EVT, etc.)
  - Nicely represented by parameters
  - BLUE, the best you can do
  - Neither underfitting nor overfitting
  - Easy to communicate

Institute and Faculty of Actuaries

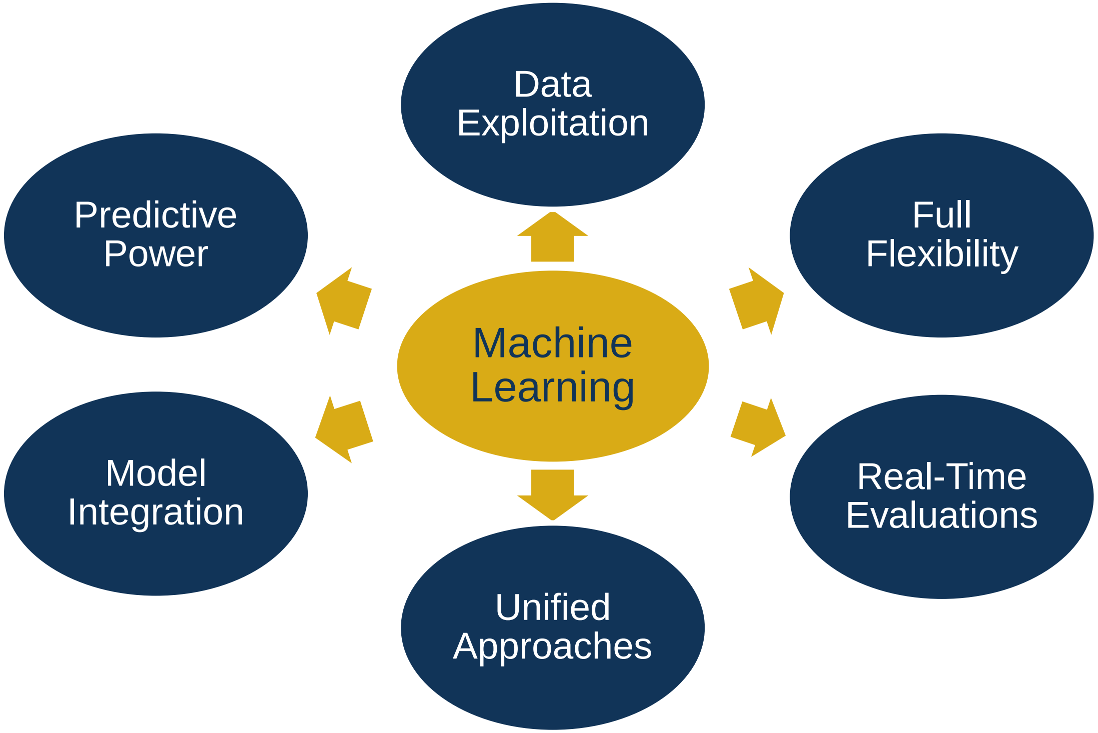
What we *experience*…

- Run-off triangles (Chain-Ladder, Fisher-Lange, Mack, etc.)
  - Easy to adapt to different branches *as long as the related portfolios are homogenous enough*
  - Easy to enhance for different purposes *as long as you accept to ignore many potential predictors*
  - Easy to visualize *as long as closing delay and accident year are the only dimensions*
  - No computational cost *as long as it is a problem (is it nowadays?!)*

- Regression methods (Gamma, Pareto, EVT, etc.)
  - Nicely represented by parameters *but non-parametric components might be significant*
  - BLUE, the best you can do *but from a variance-perspective only*
  - Neither underfitting nor overfitting *but heavy tails and clusters are poorly represented*
  - Easy to communicate *but what are you actually communicating?!*

Institute and Faculty of Actuaries

What we *need*…



Institute and Faculty of Actuaries

Institute
and Faculty
of Actuaries

Australian Bodily Injury Data: a case study

21 September 2018

The database

- Public data from the R package `CASdatasets`
- 22.036 settled personal injury insurance claims in Australia
- Accidents occurring from July 1989 to January 1999
- No zero claims reported
- Raw features:
  - Claim timeline: accident date, reporting date, closing date
  - Injury severity per injured person (up to five)
  - Legal representation (1/0)
  - Aggregated claim amount after closing
- Derived features: reporting delay, closing delay, overall injury score

Institute and Faculty of Actuaries

Assumptions

- All the information is known at reporting date
- All the claim cash-flows are paid at closing date

reporting	closing delay						
year	0	1	2	3	4	5	6
1993	**1.191**	**1.970**	**1.152**	**558**	426	277	49
1994	**874**	**1.531**	**870**	**709**	443	65	0
1995	**718**	**1.381**	**1.188**	**881**	134	0	0
1996	**529**	**1.511**	**1.314**	**174**	0	0	0
1997	585	2.020	283	0	0	0	0
1998	807	352	0	0	0	0	0
1999	14	0	0	0	0	0	0

Claim numbers

reporting	closing delay						
year	0	1	2	3	4	5	6
1993	**26.038.265**	**77.236.474**	**74.285.051**	**58.268.484**	48.169.135	38.347.205	8.444.935
1994	**16.746.725**	**42.454.730**	**42.485.823**	**55.798.479**	44.833.960	7.119.057	0
1995	**6.076.308**	**20.958.156**	**41.895.020**	**50.580.334**	8.286.925	0	0
1996	**4.090.537**	**21.665.535**	**46.736.025**	**8.158.885**	0	0	0
1997	3.787.223	33.675.352	8.543.935	0	0	0	0
1998	6.207.963	3.475.759	0	0	0	0	0
1999	36.599	0	0	0	0	0	0

Claim amounts

Institute and Faculty of Actuaries

Assumptions

- All the information is known at reporting date
- All the claim cash-flows are paid at closing date
- Training/validation records are filtered to include
  - reporting year 1993-1996
  - closing delay 0-3

reporting year	closing delay 0	1	2	3	4	5	6
1993	**1.191**	**1.970**	**1.152**	**558**	426	277	49
1994	**874**	**1.531**	**870**	**709**	443	65	0
1995	**718**	**1.381**	**1.188**	**881**	134	0	0
1996	**529**	**1.511**	**1.314**	**174**	0	0	0
1997	585	2.020	283	0	0	0	0
1998	807	352	0	0	0	0	0
1999	14	0	0	0	0	0	0

Claim numbers

reporting year	closing delay 0	1	2	3	4	5	6
1993	**26.038.265**	**77.236.474**	**74.285.051**	**58.268.484**	48.169.135	38.347.205	8.444.935
1994	**16.746.725**	**42.454.730**	**42.485.823**	**55.798.479**	44.833.960	7.119.057	0
1995	**6.076.308**	**20.958.156**	**41.895.020**	**50.580.334**	8.286.925	0	0
1996	**4.090.537**	**21.665.535**	**46.736.025**	**8.158.885**	0	0	0
1997	3.787.223	33.675.352	8.543.935	0	0	0	0
1998	6.207.963	3.475.759	0	0	0	0	0
1999	36.599	0	0	0	0	0	0

Claim amounts

Institute and Faculty of Actuaries

Assumptions

- All the information is known at reporting date
- All the claim cash-flows are paid at closing date
- Training/validation records are filtered to include
  - reporting year 1993-1996
  - closing delay 0-3
- Test records are filtered to include
  - reporting year 1997 and 1998
  - closing delay 0-2 and 0-1

reporting year	closing delay 0	1	2	3	4	5	6
1993	**1.191**	**1.970**	**1.152**	**558**	426	277	49
1994	**874**	**1.531**	**870**	**709**	443	65	0
1995	**718**	**1.381**	**1.188**	**881**	134	0	0
1996	**529**	**1.511**	**1.314**	**174**	0	0	0
1997	585	2.020	283	0	0	0	0
1998	807	352	0	0	0	0	0
1999	14	0	0	0	0	0	0

Claim numbers

reporting year	closing delay 0	1	2	3	4	5	6
1993	**26.038.265**	**77.236.474**	**74.285.051**	**58.268.484**	48.169.135	38.347.205	8.444.935
1994	**16.746.725**	**42.454.730**	**42.485.823**	**55.798.479**	44.833.960	7.119.057	0
1995	**6.076.308**	**20.958.156**	**41.895.020**	**50.580.334**	8.286.925	0	0
1996	**4.090.537**	**21.665.535**	**46.736.025**	**8.158.885**	0	0	0
1997	3.787.223	33.675.352	8.543.935	0	0	0	0
1998	6.207.963	3.475.759	0	0	0	0	0
1999	36.599	0	0	0	0	0	0

Claim amounts

Institute and Faculty of Actuaries

Assumptions

- All the information is known at reporting date
- All the claim cash-flows are paid at closing date
- Training/validation records are filtered to include
  - reporting year 1993-1996
  - closing delay 0-3
- Test records are filtered to include
  - reporting year 1997 and 1998
  - closing delay 0-2 and 0-1
- Remaining data are ignored for modelling reasons

reporting	closing delay						
year	0	1	2	3	4	5	6
1993	1.191	1.970	1.152	558	426	277	49
1994	874	1.531	870	709	443	65	0
1995	718	1.381	1.188	881	134	0	0
1996	529	1.511	1.314	174	0	0	0
1997	585	2.020	283	0	0	0	0
1998	807	352	0	0	0	0	0
1999	14	0	0	0	0	0	0

Claim numbers

reporting	closing delay						
year	0	1	2	3	4	5	6
1993	26.038.265	77.236.474	74.285.051	58.268.484	48.169.135	38.347.205	8.444.935
1994	16.746.725	42.454.730	42.485.823	55.798.479	44.833.960	7.119.057	0
1995	6.076.308	20.958.156	41.895.020	50.580.334	8.286.925	0	0
1996	4.090.537	21.665.535	46.736.025	8.158.885	0	0	0
1997	3.787.223	33.675.352	8.543.935	0	0	0	0
1998	6.207.963	3.475.759	0	0	0	0	0
1999	36.599	0	0	0	0	0	0

Claim amounts

Institute and Faculty of Actuaries

Institute and Faculty of Actuaries

Closing Delay Estimation

21 September 2018

How <u>long</u> will it take to pay the claim amount?

<u>Categorical</u> target variable (0-3 years)

- *Multinomial Regression* → based on a GLM with the canonical link function for multinomial-distributed variables
- *Naïve Bayes* → based on the Bayes Theorem assuming independence among predictors
- *Nearest Neighbours* → based on the k nearest records according to the Euclidean distance
- *Classification Tree* → based on recursive partitioning for training and backward pruning for validation

Institute and Faculty of Actuaries

Multinomial Regression

$$\hat{y}_i = E[y_i] = g_M^{-1}\left(\beta_0 + \sum_j \beta_j x_{ij}\right)$$

Actual	*Predicted Class*				*Cases*	*Errors*	*Errors*
Class	0	1	2	3	*Number*	*Number*	*Percentage*
0	97	1.805	73	12	1.987	1.890	95,12%
1	130	3.339	329	61	3.859	520	13,47%
2	83	2.082	445	71	2.681	2.236	83,40%
3	44	1.027	207	126	1.404	1.278	91,03%
Total					**9.931**	**5.924**	**59,65%**

Multinomial regression summary results using training data

Actual	*Predicted Class*				*Cases*	*Errors*	*Errors*
Class	0	1	2	3	*Number*	*Number*	*Percentage*
0	72	1.194	44	15	1.325	1.253	94,57%
1	92	2.185	220	37	2.534	349	13,77%
2	52	1.444	287	60	1.843	1.556	84,43%
3	29	664	155	70	918	848	92,37%
Total					**6.620**	**4.006**	**60,51%**

Multinomial regression summary results using validation data

Institute and Faculty of Actuaries

Naïve Bayes

$$\hat{p}_i(\Delta_i = k) = \frac{P(\Delta_i = k)\prod_{j=1}^{n} P(x_j = x_{ij}|\Delta_i = k)}{\sum_{h=0}^{3} P(\Delta_i = h)\prod_{j=1}^{n} P(x_j = x_{ij}|\Delta_i = h)}$$

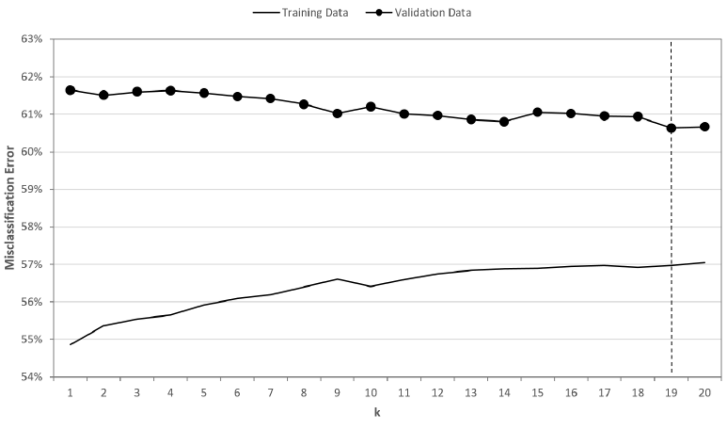
Actual	*Predicted Class*				*Cases*	*Errors*	*Errors*
Class	0	1	2	3	*Number*	*Number*	*Percentage*
0	740	1.069	151	27	1.987	1.247	62,76%
1	870	2.210	689	90	3.859	1.649	42,73%
2	434	1.380	764	103	2.681	1.917	71,50%
3	242	730	281	151	1.404	1.253	89,25%
Total					**9.931**	**6.066**	**61,08%**

Naive Bayes summary results using training data

Actual	*Predicted Class*				*Cases*	*Errors*	*Errors*
Class	0	1	2	3	*Number*	*Number*	*Percentage*
0	477	708	122	18	1.325	848	64,00%
1	536	1.454	490	54	2.534	1.080	42,62%
2	284	988	483	88	1.843	1.360	73,79%
3	153	445	226	94	918	824	89,76%
Total					**6.620**	**4.112**	**62,11%**

Naive Bayes summary results using validation data

Nearest Neighbours



Actual Class	*Predicted Class* 0	1	2	3	*Cases Number*	*Errors Number*	*Errors Percentage*
0	472	1.389	105	21	1.987	1.515	76,25%
1	367	3.140	286	66	3.859	719	18,63%
2	222	1.869	502	88	2.681	2.179	81,28%
3	115	938	191	160	1.404	1.244	88,60%
Total					**9.931**	**5.657**	**56,96%**

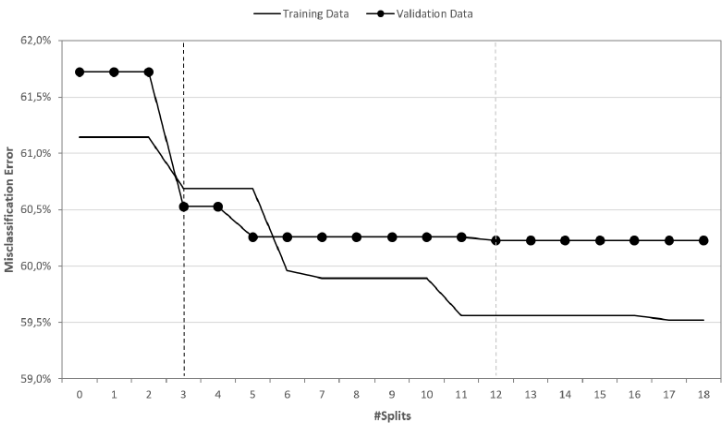
Nearest neighbours summary results using training data

Actual Class	*Predicted Class* 0	1	2	3	*Cases Number*	*Errors Number*	*Errors Percentage*
0	267	959	83	16	1.325	1.058	79,85%
1	269	1.990	224	51	2.534	544	21,47%
2	164	1.355	248	76	1.843	1.595	86,54%
3	75	614	128	101	918	817	89,00%
Total					**6.620**	**4.014**	**60,63%**

Nearest neighbours summary results using validation data

Institute and Faculty of Actuaries

Classification Tree



Actual	*Predicted Class*				*Cases*	*Errors*	*Errors*
Class	0	1	2	3	*Number*	*Number*	*Percentage*
0	0	1.892	57	38	1.987	1.987	100,00%
1	0	3.412	258	189	3.859	447	11,58%
2	0	2.101	347	233	2.681	2.334	87,06%
3	0	1.094	49	261	1.404	1.143	81,41%
Total					**9.931**	**5.911**	**59,52%**

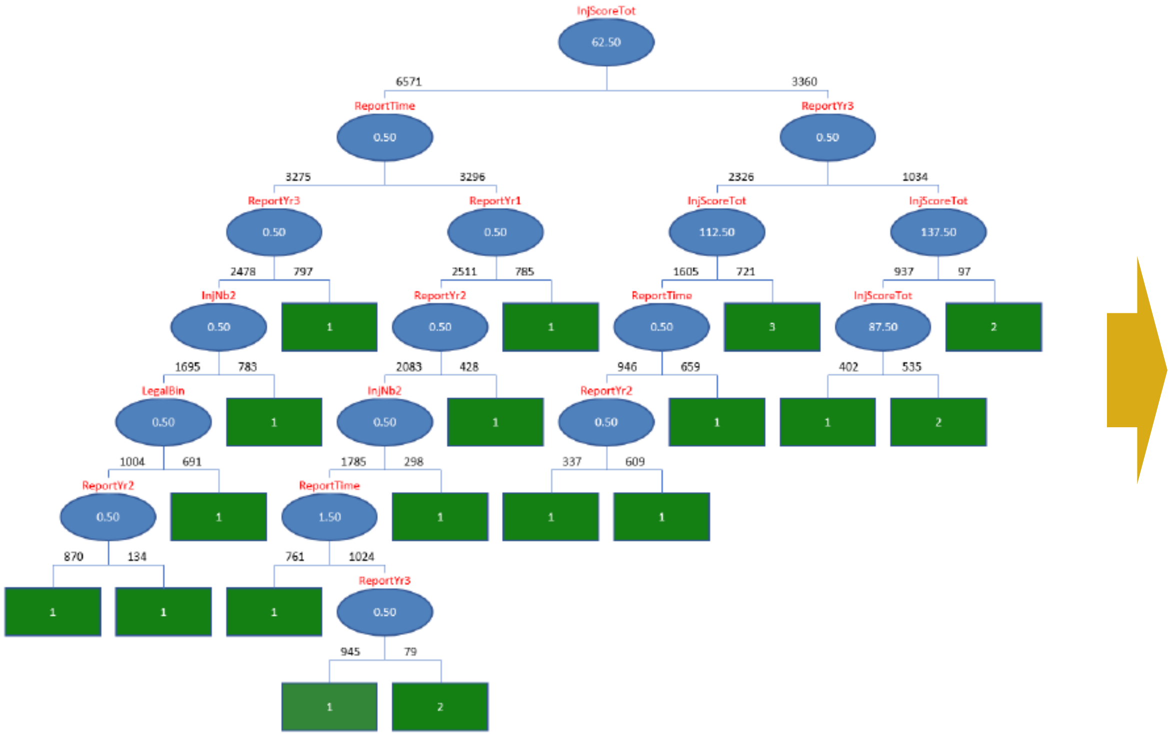
Classification tree summary results using training data

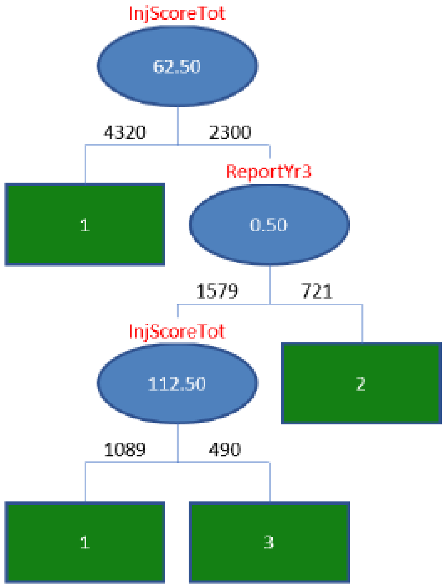
Actual	*Predicted Class*				*Cases*	*Errors*	*Errors*
Class	0	1	2	3	*Number*	*Number*	*Percentage*
0	0	1.229	67	29	1.325	1.325	100,00%
1	0	2.121	294	119	2.534	413	16,30%
2	0	1.363	315	165	1.843	1.528	82,91%
3	0	696	45	177	918	741	80,72%
Total					**6.620**	**4.007**	**60,53%**

Classification tree summary results using validation data

Institute and Faculty of Actuaries

Classification Tree





Institute and Faculty of Actuaries

Institute
and Faculty
of Actuaries

Claim Amount Prediction

21 September 2018

How <u>much</u> will it take to pay the claim amount?

<u>Numerical</u> target variable ($ amount)

Gamma Regression	→	based on a GLM with the canonical link function for gamma-distributed variables
Regression Tree	→	based on recursive partitioning for training and backward pruning for validation
Neural Network	→	based on back propagation and arctangent as activation function

Institute and Faculty of Actuaries

Gamma Regression

$$\hat{y}_i = E[y_i] = g_\Gamma^{-1}\left(\beta_0 + \sum_j \beta_j x_{ij}\right)$$

	#Coeffs	R^2	Adj. R^2	Intercept	InjScoreTot	InjNb2	InjNb3	InjNb4	InjNb5	LegalBin	ReportTime	ReportYr1	ReportYr2	ReportYr3	FinTime1	FinTime2	FinTime3
Stepwise Selection	1	0,0%	0,0%	x													
	2	7,9%	7,9%	x													x
	3	16,3%	16,3%	x							x						x
	4	25,8%	25,8%	x							x					x	x
	5	30,8%	30,8%	x							x				x	x	x
	6	34,7%	34,7%	x	x						x				x	x	x
	7	36,8%	36,8%	x	x					x	x				x	x	x
	8	38,0%	38,0%	x	x					x	x			x	x	x	x
	9	**39,4%**	**39,3%**	**x**	**x**					**x**	**x**		**x**	**x**	**x**	**x**	**x**

		Intercept	InjScoreTot	InjNb2	InjNb3	InjNb4	InjNb5	LegalBin	ReportTime	ReportYr1	ReportYr2	ReportYr3	FinTime1	FinTime2	FinTime3
Gamma Regression	*Coefficient*	8,5839	0,0084					0,3353	0,3985		-0,4863	-0,5905	0,5660	1,2402	1,6610
	Standard Error	0,0478	0,0005					0,0342	0,0151		0,0421	0,0453	0,0444	0,0484	0,0572
	t-Statistic	179,54	18,31					9,81	26,40		-11,56	-13,02	12,76	25,61	29,04
	P-Value	0,0000	0,0000					0,0000	0,0000		0,0000	0,0000	0,0000	0,0000	0,0000
	Conf. Interval Lower	8,4849	0,0075					0,2677	0,3665		-0,5686	-0,6807	0,4782	1,1439	1,5476
	Conf. Interval Upper	8,6840	0,0093					0,4027	0,4308		-0,4032	-0,4996	0,6532	1,3362	1,7752

Institute and Faculty of Actuaries

Regression Tree

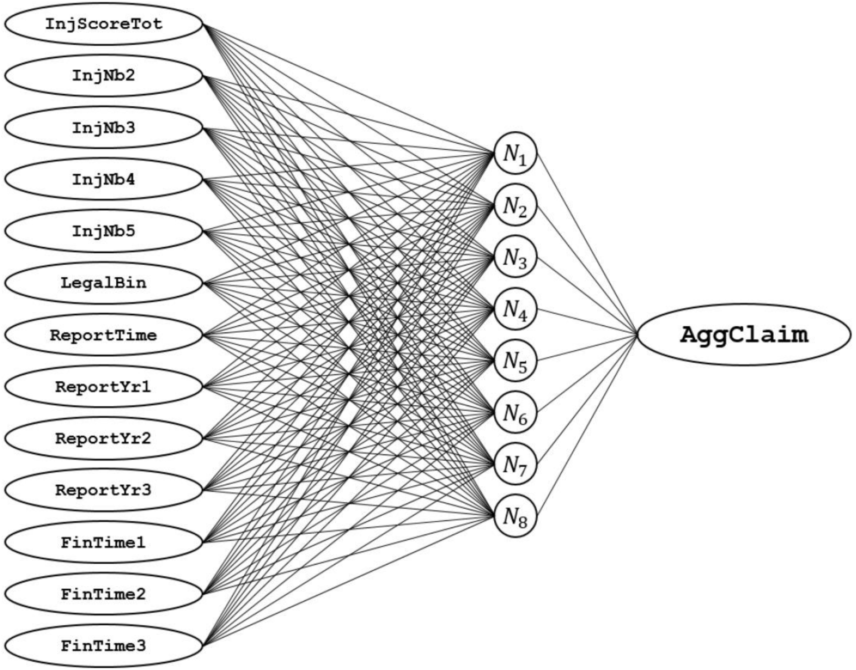
minimum records per leaf	average training error	average validation error	average overall error	percentage weighted error	percentage overall error
10	28.262	28.554	28.379	16,63%	0,38%
20	28.393	28.622	28.485	16,74%	0,38%
30	27.819	28.020	27.899	14,28%	0,39%
40	27.819	28.020	27.899	14,28%	0,39%
50	**26.695**	**27.528**	**27.028**	**3,67%**	**0,29%**
60	26.982	27.477	27.180	5,06%	0,22%
70	27.088	27.326	27.183	5,54%	0,10%
80	27.154	27.393	27.250	5,49%	0,09%
90	27.185	27.331	27.243	6,10%	0,01%
100	27.273	27.385	27.318	6,30%	0,15%
110	27.339	27.457	27.386	6,83%	0,16%
120	27.525	27.568	27.542	9,14%	0,16%
130	27.603	27.596	27.600	9,47%	0,10%
140	27.647	27.679	27.660	9,83%	0,25%
150	27.729	27.753	27.739	9,85%	0,41%
160	27.632	27.680	27.651	8,43%	0,37%
170	27.652	27.662	27.656	8,45%	0,40%
180	27.743	27.718	27.733	9,09%	0,39%
190	27.764	27.741	27.755	9,38%	0,43%
200	27.806	27.797	27.802	9,60%	0,51%
210	27.830	27.777	27.809	9,59%	0,48%
220	27.842	27.788	27.820	9,81%	0,50%
230	27.826	27.766	27.802	9,67%	0,48%
240	27.952	27.866	27.918	9,93%	0,54%
250	27.971	27.906	27.945	10,18%	0,55%
260	27.971	27.913	27.948	10,96%	0,56%
270	28.000	28.116	28.046	11,02%	0,67%
280	28.012	28.139	28.063	11,30%	0,72%
290	28.032	28.171	28.088	11,50%	0,69%
300	28.196	28.430	28.289	13,86%	0,63%

Level	Item ID	Item Type	Parent ID	Left Child ID	Right Child ID	Split Variable	Split Value	Cases	Prediction
0	0	Node	-	1	2	InjScoreTot	137,5	6620	36.078
1	1	Node	0	3	4	ReportTime	1,5	6420	33.046
1	2	Node	0	5	6	ReportTime	1,5	200	134.457
2	3	Node	1	7	8	FinTime3	0,5	5032	24.782
2	4	Node	1	9	10	FinTime3	0,5	1388	61.671
2	5	Node	2	11	12	FinTime3	0,5	154	103.858
2	6	Node	2	13	14	ReportYr3	0,5	46	243.674
3	7	Node	3	15	16	FinTime2	0,5	4340	19.989
3	8	Node	3	17	18	ReportTime	0,5	692	54.986
3	9	Node	4	19	20	InjScoreTot	87,5	1226	54.457
3	10	Node	4	21	22	InjScoreTot	62,5	162	112.333
3	11	Node	5	23	24	FinTime2	0,5	101	78.539
3	12	Node	5	25	26	ReportTime	0,5	53	144.538
3	13	Node	6	27	28	ReportYr2	0,5	43	266.160
3	14	Leaf	6	-	-	-	-	3	83.462
4	15	Node	7	29	30	ReportTime	0,5	2898	13.545
4	16	Node	7	31	32	ReportTime	0,5	1442	33.727
4	17	Node	8	33	34	InjScoreTot	62,5	523	49.081
4	18	Node	8	35	36	ReportYr2	0,5	169	72.337
4	19	Node	9	37	38	FinTime2	0,5	1097	49.800
4	20	Node	9	39	40	ReportYr3	0,5	129	100.949
4	21	Node	10	41	42	InjScoreTot	37,5	116	88.581
4	22	Node	10	43	44	InjNb2	0,5	46	158.795
4	23	Leaf	11	-	-	-	-	38	46.638
4	24	Node	11	45	46	LegalBin	0,5	63	89.273
4	25	Node	12	47	48	LegalBin	0,5	43	133.284
4	26	Leaf	12	-	-	-	-	10	195.888
4	27	Node	13	49	50	ReportYr1	0,5	39	286.621
4	28	Leaf	13	-	-	-	-	4	92.235

Institute and Faculty of Actuaries

Neural Network

number of hidden neurons	average training error	average validation error	average overall error	percentage weighted error	percentage overall error
1	30.420	30.068	30.279	16,23%	12,50%
2	29.634	29.621	29.629	15,20%	9,17%
3	29.706	29.838	29.758	13,96%	12,09%
4	28.858	28.947	28.894	8,96%	8,57%
5	29.028	29.285	29.131	11,26%	9,27%
6	28.734	28.828	28.771	10,08%	8,52%
7	28.848	29.338	29.044	11,13%	9,80%
8	**28.068**	**28.607**	**28.284**	**8,30%**	**5,57%**
9	28.609	29.273	28.875	8,89%	8,28%
10	28.615	29.411	28.933	10,66%	6,89%
11	28.318	29.056	28.613	10,04%	8,81%
12	28.553	29.229	28.824	10,60%	8,02%
13	28.376	28.975	28.616	9,81%	8,33%
14	28.375	29.368	28.772	9,82%	7,10%
15	28.261	28.900	28.516	8,82%	6,44%
16	28.480	29.020	28.696	9,64%	7,19%
17	28.402	29.299	28.761	10,07%	7,77%
18	28.188	28.939	28.488	9,00%	6,90%
19	28.321	28.831	28.525	8,64%	6,55%
20	28.044	28.890	28.382	9,06%	5,76%



Institute and Faculty of Actuaries

Results

reporting year	closing delay 0	1	2	3
1993	26.038.265	77.236.474	74.285.051	58.268.484
1994	16.746.725	42.454.730	42.485.823	55.798.479
1995	6.076.308	20.958.156	41.895.020	50.580.334
1996	4.090.537	21.665.535	46.736.025	8.158.885

Actual data

reporting year	closing delay 0	1	2	3
1993	23.942.277	72.096.426	85.372.988	68.135.627
1994	15.615.349	43.607.867	49.123.608	69.977.076
1995	6.294.610	21.590.823	41.199.048	51.380.998
1996	4.338.858	23.040.826	46.178.322	10.158.513

Gamma regression

reporting year	closing delay 0	1	2	3
1993	27.777.729	74.715.426	74.860.543	53.671.653
1994	17.317.427	44.645.051	41.115.866	56.014.481
1995	7.819.221	22.443.511	40.888.916	51.092.375
1996	5.285.040	22.293.701	46.183.852	9.048.303

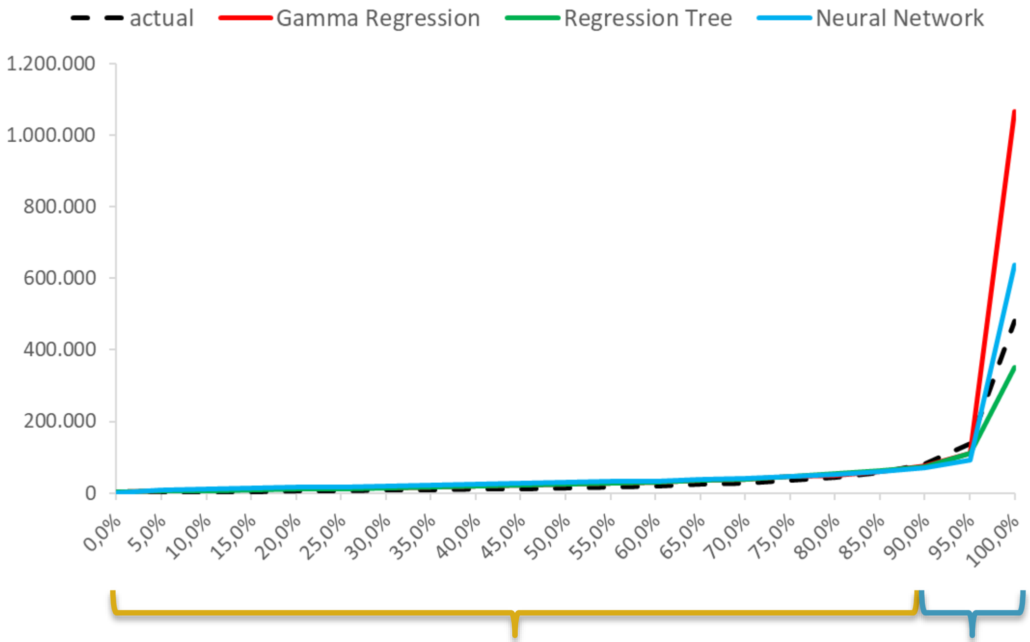
Regression tree

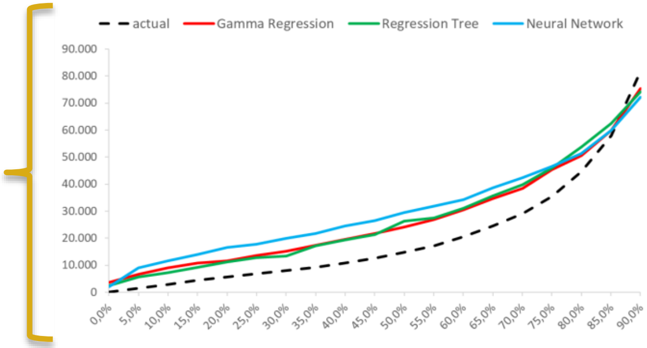
reporting year	closing delay 0	1	2	3
1993	30.323.454	80.122.598	74.090.121	56.198.161
1994	21.876.621	49.791.426	44.957.040	51.693.100
1995	8.007.124	25.605.931	47.106.288	49.418.199
1996	5.358.636	27.349.566	46.177.586	8.455.383

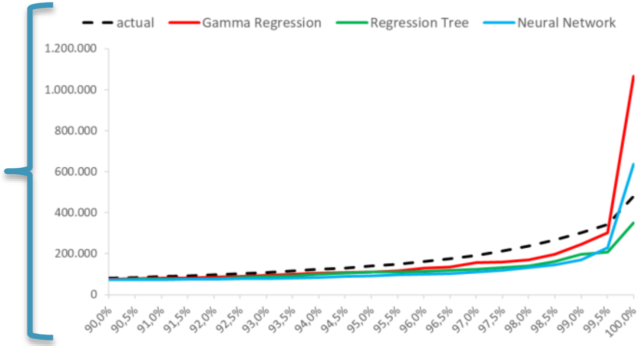
Neural network

Institute and Faculty of Actuaries

Results







Institute and Faculty of Actuaries

Institute and Faculty of Actuaries

Reserve Evaluation

21 September 2018

The model

$$\hat{R}_i = \sum_{k=0}^{3} \overbrace{\hat{p}_i(\Delta_i = k)}\underbrace{\hat{C}_{\Delta_i = k}}$$

Multinomial Regression

Naïve Bayes

Nearest Neighbours

Classification Tree

Gamma Regression

Regression Tree

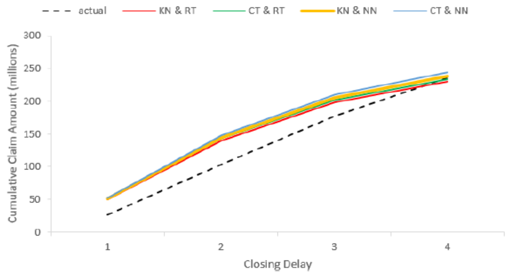
Neural Network

Institute
and Faculty
of Actuaries

Results – reporting years 1993 and 1994

Actual claim amount 235.828.275	*gamma regression*	*regression tree*	*neural network*
multinomial regression	252.028.878	230.418.987	241.288.489
naive Bayes	239.075.514	217.568.447	231.965.733
nearest neighbours	249.685.356	229.945.901	238.026.479
classification tree	258.083.300	233.422.075	243.846.286

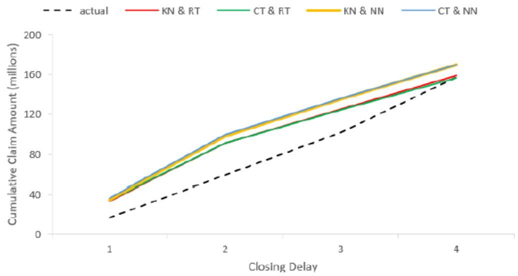
Δ%	*gamma regression*	*regression tree*	*neural network*
multinomial regression	6,87%	-2,29%	2,32%
naive Bayes	1,38%	-7,74%	-1,64%
nearest neighbours	5,88%	-2,49%	0,93%
classification tree	9,44%	-1,02%	3,40%



Predictive performance for reporting year 1993

Actual claim amount 157.485.757	*gamma regression*	*regression tree*	*neural network*
multinomial regression	186.011.183	159.763.704	172.850.776
naive Bayes	177.763.171	149.619.919	165.548.345
nearest neighbours	182.202.506	159.124.199	169.743.174
classification tree	182.470.513	156.827.235	170.113.669

Δ%	*gamma regression*	*regression tree*	*neural network*
multinomial regression	18,11%	1,45%	9,76%
naive Bayes	12,88%	-4,99%	5,12%
nearest neighbours	15,69%	1,04%	7,78%
classification tree	15,86%	-0,42%	8,02%



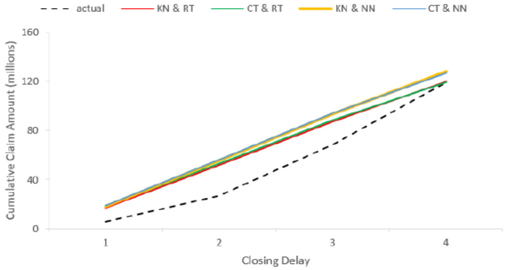
Predictive performance for reporting year 1994

Institute and Faculty of Actuaries

Results – reporting years 1995 and 1996

Actual claim amount 119.509.819	*gamma regression*	*regression tree*	*neural network*
multinomial regression	118.524.978	122.550.580	129.893.618
naive Bayes	126.742.857	129.321.735	136.983.874
nearest neighbours	117.758.364	120.167.319	128.321.910
classification tree	116.673.918	119.593.669	126.714.629

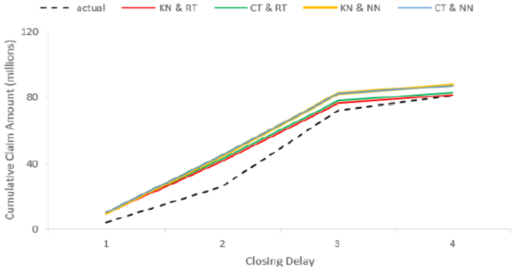
Δ%	*gamma regression*	*regression tree*	*neural network*
multinomial regression	-0,82%	2,54%	8,69%
naive Bayes	6,05%	8,21%	14,62%
nearest neighbours	-1,47%	0,55%	7,37%
classification tree	-2,37%	0,07%	6,03%



Predictive performance for reporting year 1995

Actual claim amount 80.650.982	*gamma regression*	*regression tree*	*neural network*
multinomial regression	82.330.316	81.844.353	87.189.365
naive Bayes	87.835.102	86.321.719	91.700.340
nearest neighbours	83.617.031	81.500.277	87.200.645
classification tree	83.901.388	82.788.649	87.006.662

Δ%	*gamma regression*	*regression tree*	*neural network*
multinomial regression	2,08%	1,48%	8,11%
naive Bayes	8,91%	7,03%	13,70%
nearest neighbours	3,68%	1,05%	8,12%
classification tree	4,03%	2,65%	7,88%



Predictive performance for reporting year 1996

Institute and Faculty of Actuaries

Results – overall

Actual claim amount 593.474.833	*gamma regression*	*regression tree*	*neural network*
multinomial regression	638.895.354	594.577.625	631.222.248
naive Bayes	631.416.644	582.831.820	626.198.292
nearest neighbours	633.263.257	590.737.696	623.292.207
classification tree	641.129.119	592.631.628	627.681.248

Δ%	*gamma regression*	*regression tree*	*neural network*
multinomial regression	7,65%	0,19%	6,36%
naive Bayes	6,39%	-1,79%	5,51%
nearest neighbours	6,70%	-0,46%	5,02%
classification tree	8,03%	-0,14%	5,76%

Overall predictive performance

Institute and Faculty of Actuaries

Limitations

- Because of limited data, we could not compare our individual claim reserve estimation to triangle-based reserve estimations, although this would have been of scarce usefulness as our data shows too little homogeneity.
- Most of machine learning methods do not allow for estimation of data that is not well represented in the dataset, for instance, claims paid later than three years.
- The reserve estimated in this study is the RBNYS only, although the IBNYR reserve might be material for some branches. Our data is insufficient to estimate the latter, but the model could be easily adapted for that as long as right data is provided.

Institute and Faculty of Actuaries

Conclusions

- While the closing delay estimation is just slightly better than that based on prior probabilities regardless of the machine learning technique used, the claim amount prediction is not only relevant but even significantly better than that from the GLM based on the gamma distribution and its canonical link function (which is a very common choice in the actuarial practice).
- For sake of brevity, part of our analysis was not included in this presentation, that is, the results on the test dataset (reporting years 1997-1998) as well as the prediction of zero-claims and large claims.
- The references to the full study are reported in the next slide.

Institute and Faculty of Actuaries

Institute and Faculty of Actuaries

References

21 September 2018

References

- M. Aleandri, "*Case Reserving in Non-Life Practice using Individual Data and Machine Learning*", Rapporto Tecnico, Univ. "La Sapienza", Rome, 2017
- M. Aleandri, "Data Science and Machine Learning in Actuarial Practice", SA0 Research Dissertation (*under review*), Institute and Faculty of Actuaries, London, 2018

Institute and Faculty of Actuaries

Questions

Comments

The views expressed in this presentation are those of invited contributors and not necessarily those of the IFoA. The IFoA do not endorse any of the views stated, nor any claims or representations made in this presentation and accept no responsibility or liability to any person for loss or damage suffered as a consequence of their placing reliance upon any view, claim or representation made in this presentation.

The information and expressions of opinion contained in this publication are not intended to be a comprehensive study, nor to provide actuarial advice or advice of any nature and should not be treated as a substitute for specific advice concerning individual situations. On no account may any part of this presentation be reproduced without the written permission of the IFoA.

Institute and Faculty of Actuaries