PREDICTION OF CLAIMS IN EXPORT CREDIT FINANCE: A COMPARISON OF FOUR MACHINE LEARNING TECHNIQUES

Mathias Bärtl

Offenburg University, Germany

Simone Krummaker

The Business School (formerly Cass), London, UK

IFTI GLOBAL SYMPOSIUM 2020

MOTIVATION

- Relevance of claims prediction
- ML approaches to predict claims
- Do ML approaches add value compared to simpler techniques?
- Very little research on export credit insurance claims

AGENDA

01 Introduction and Motivation

02 Background on Export Credit Insurance and Claims Prediction

03 Introduction to Machine Learning Approaches and Dataset

04 Results and Discussion

05 Conclusion and Outlook

BACKGROUND

Claims prediction is a critical process for insurers

- Claims arrive as a stochastic process
  - Uncertain number of claims
  - Uncertain amount of claims

- Premiums are fixes and paid upfront, before total amount of business expenses and claims is known

- Claims in export credit agencies are a rare event but are significantly influenced by the local and global economic context
  - Geo-political events
  - Sars-CoV-2 pandemic event on supply chains, business continuity, insolvencies, …

BACKGROUND

- ML approaches can be well suited to provide more accurate claims predictions
- Current developments in the global trade and export environment might lead to a more volatile claims situation now and in the near future
- While insurers appreciate automatization and further input into decision making, they are still keen to apply common sense/human judgement

AIM OF THE PAPER

- to assess the performance of ML techniques in identifying export credit insurance claims
- to assess the potential performance loss when tested under near- realistic forecasting condition
- to evaluate their performance against a simple benchmark (BM) technique

AGENDA

01 Introduction and Motivation

02 Background on Export Credit Insurance and Claims Prediction

03 Introduction to Machine Learning Approaches and Dataset

04 Results and Discussion

05 Conclusion and Outlook

TRADE AND EXPORT CREDIT INSURANCE

- Covers the loss of receivables due to the risk of non-payment of the buyer
  - Full or part default
  - protracted default (extended late payment)
  - due to commercial risks (e.g. insolvency of the buyer)
- Pre-shipment cover can usually also be obtained
  - Covers against risks of contract frustration during the manufacturing period

➢ Protects exporter’s cash flows and unwanted cash flow volatility due to unsystematic risk

ECI MARKET

Private Insurance Companies

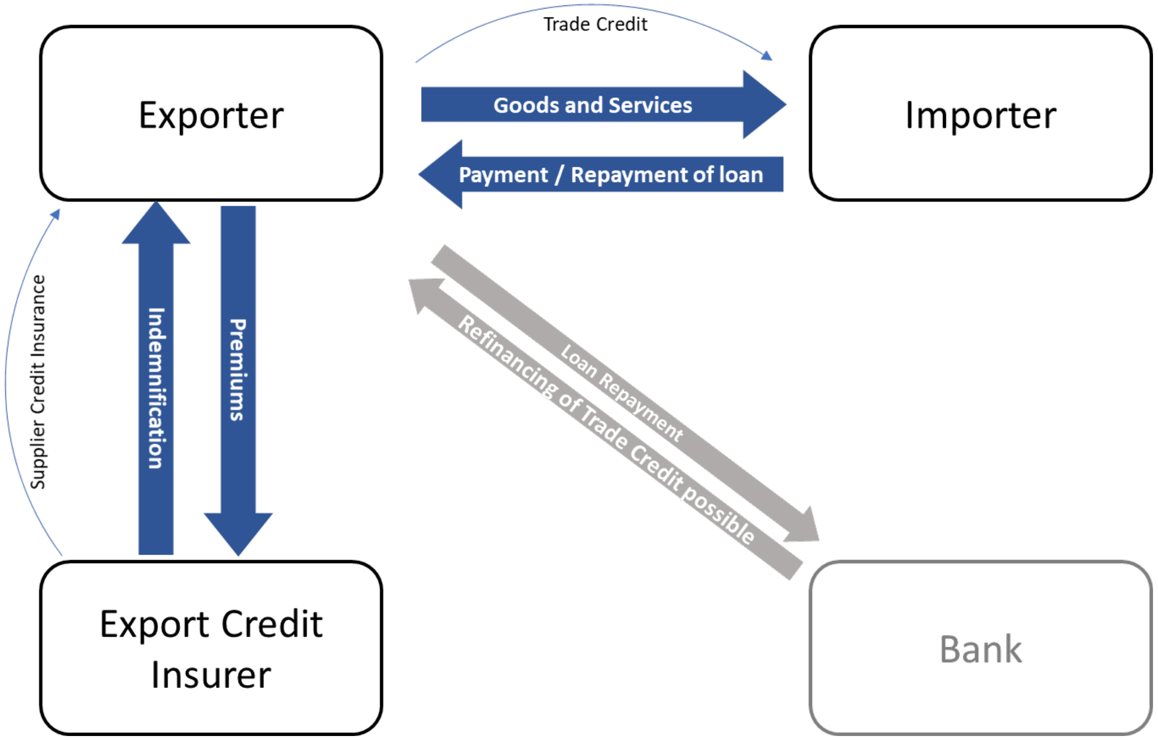
- Usually are only happy to accept short-term export credit risks
  - of up to 2 years and under “normal conditions”
  - Usually do not cover medium- or long-term credit risks
- Marketable Risks
  - Risks for which a private insurance market in principle exists
  - That might also meant that the private insurers are able to obtain sufficient reinsurance capacity
- High market concentration: Oligopoly (with keen competition)
  - The big 3: EulerHermes 26%, Atradius 15%, Coface 15% of global credit insurance
  - Plus Sinosure 22% (covers 90% of all Chinese exports)

Export Credit Agencies (ECAs)

- ECAs are the largest source of government funding for private businesses
  - Official branches of the government, public/government backed providers
- Offer export credit insurance, guarantees and sometimes financing for **non-marketable risks**
  - Mainly medium- and long-term risks
- Purpose
  - ECAs are integral to government trade and foreign aid strategies
  - to foster international trade
  - To promote exports (and by this to contribute to employment and economic growth)

EXAMPLE: SUPPLIER CREDIT COVER

- most common form of short- and mid-term trade finance
- exporter gives the buyer time for payment after the delivery
- supplier bears the risk of default which can be covered with a supplier credit insurance
- supplier/exporter might refinance the trade credit granted to the buyer using e.g banks
- then, the indemnification may be assigned to the refinancing organisation in the case of default of the buyer.



CLAIMS PREDICTION AND RESERVING

- Stochastic and deterministic methods are well established (e.g. Baudry and Robert 2019)
  - Chain ladder
  - Bornhuetter-Ferguson method
- Market and regulatory developments ask for more sophisticated methods (England and Verrall 2002; Verrall et al. 2012)
- ML methods are beneficial for claims prediction (Wüthrich 2018a, 2018b, Thesmar et al. 2019)

AGENDA

01 Introduction and Motivation

02 Background on Export Credit Insurance and Claims Prediction

03 Introduction to Machine Learning Approaches and Dataset

04 Results and Discussion

05 Conclusion and Outlook

SUPERVISED MACHINE LEARNING

- Purpose
  - to uncover potential relationships between independent and one or several dependent variables
  - to find a function that allows a good prediction of a target attribute, based on available input attributes

- literature provides a wide range of ML applications
  - including Naïve Bayesian Classifiers, Bayesian Networks, Logistic Regression, Decision Trees (DT), Conditional Inference Trees, Random Forests (RF), Support Vector Machines, k-Nearest-Neighbour and Neuronal Networks (NN); The Least Absolute Shrinkage and Selection Operator (LASSO) algorithm is used occasionally in economic applications and is alleged to be most familiar to economists (Mullainathan and Spiess 2017)
  - All these techniques are, in principle, suitable in supporting the prediction of claims

- No widely accepted approach to determine which (type of) problem is best addressed with which ML technique (Kuhn and Johnson 2013; Wanke and Barros 2016)
  - Field of application is key to determine, compare and judge the performance of different algorithms (Singh et al. 2016)
  - Popular to apply several techniques and compare (e.g. Fauzan and Murfi 2018; Lorena et al. 2011; Mullainathan and Spiess 2017; Razi and Athappilli 2005; Singh et al. 2016; Weerasinghe and Wijegunasekara 2016)

- We are comparatively investigating Decision Trees (DT), Random Forests (RF), Neuronal Networks (NN) and Probabilistic Neuronal Networks (PNN) to predict claims in export credit insurance
  Detailed description of each method in the paper

DATA

Provided by Berne Union

- Covers 85 Export Credit Agencies, private insurers and multilateral institutions from 73 countries
- Account for 13% of the global cross-border trade (2019: USD 2.5 trillion cover volume, USD 6 billion claims paid) | MLT business: 83% covered by ECAs (2018)
- Best overall proxy for trade credit in general

Data set

- Period of 2005 to 2018
- MLT ECA business
- Variables include ECA, destination country, activity (insurance or lending), volume of new commitments by type (sovereign, Other Public, Banks, Corporates and Projects), the volume of claims and recoveries (political, commercial, total), offers, reinsurance, exposure, staff, premium income, administrative costs and cash flow)

DESCRIPTIVE STATISTICS

Year [1]	Number of Records	Exposure		New Commitments		Claims Paid	
		Mean	SD	Mean	SD	Mean	SD
2007	1983	254.37	785.74	72.24	350.92	0.58	4.13
2008	2028	248.34	804.19	73.51	334.62	0.49	4.48
2009	2094	278.96	927.32	91.29	528.51	1.44	27.37
2010	2063	284.32	873.14	82.59	343.57	0.82	6.37
2011	2072	288.09	876.39	86.18	364.00	1.07	10.15
2012	2078	303.35	897.73	79.91	324.73	1.02	11.36
2013	2061	320.35	939.42	71.49	275.30	1.08	9.69
2014	2150	296.78	883.90	70.46	356.33	0.93	10.03
2015	2194	301.25	901.09	64.78	347.95	1.38	24.78
2016	2189	308.82	971.67	58.51	330.23	1.34	13.06
2017	2239	306.62	985.34	57.85	374.20	1.18	9.42
2018	2245	301.31	1007.71	59.29	314.81	1.40	12.28

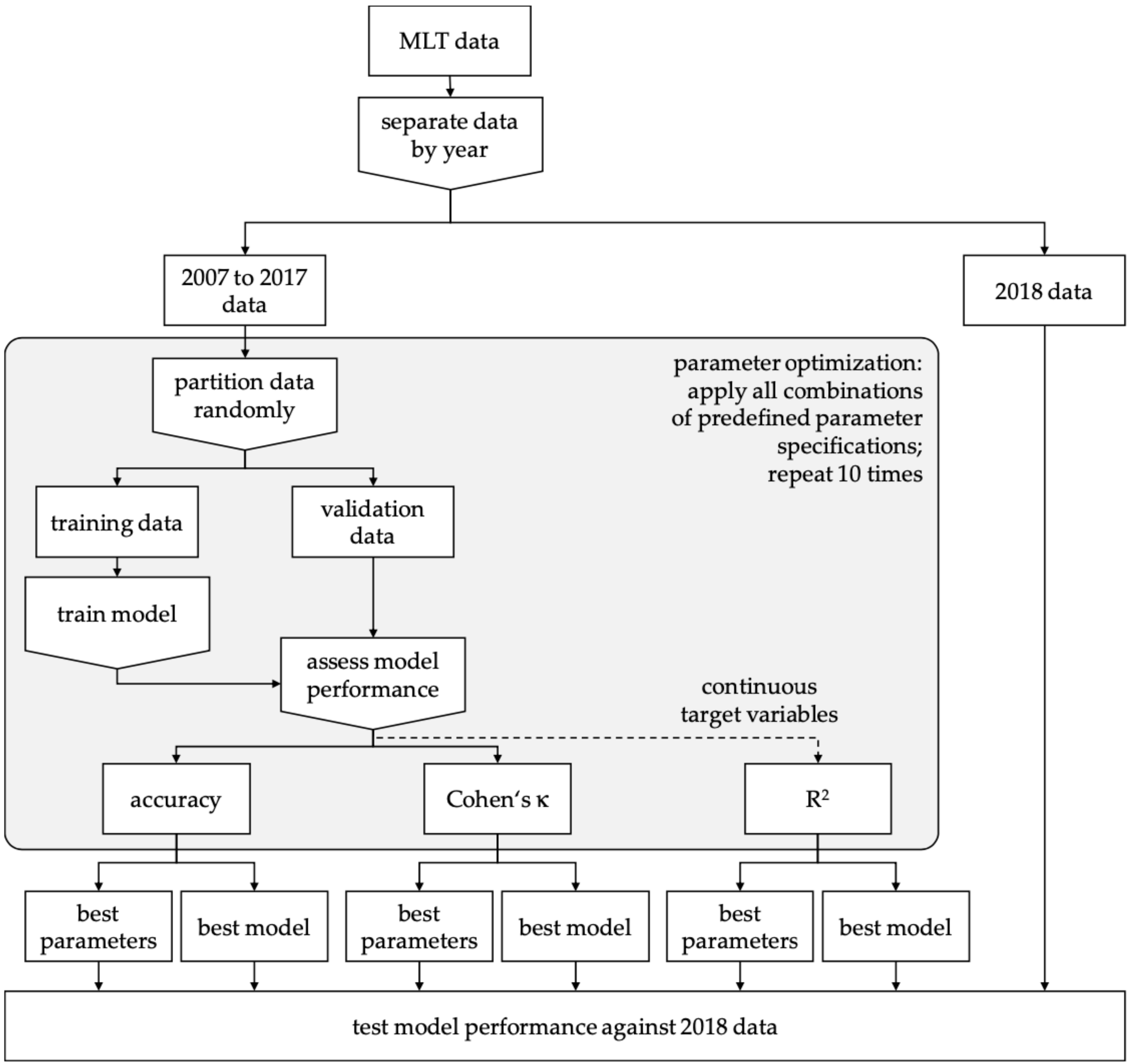
[1] Data was enriched to include simple trend estimates based on the current and two antecedent years (see Appendix A for details). Records from 2005 and 2006 could therefore not be used in support of the actual ML exercise.

METHODOLOGY

- Objective: to compare the performance of the different ML techniques, we train models to solve prediction tasks with different degrees of difficulty
  - **“Claims YES/NO”:** predict whether or not a given export finance condition will incur claims
  - **“Claim ratio class”:** magnitude of claims, expressed as five classes of claims/exposure-ratios.
  - **“Claim ratio”:** actual claim ratio, measured in terms of claims/exposure.
- Software: KNIME
- Training, Validation and Test Data
  - records from 2007 to 2017 are used for training and validation
  - 2018 data are used to test the predictions

GENERAL MODELLING CONSIDERATIONS

- ECA business and risk is shaped by the
  - size of its national economy and export profile
  - Political, judicial and commercial structure and stability of the destination countries
- Training-Validation Gateway considerations
  - We prevent ML algorithms from knowing agents of export/financing transactions
  - Only generic information (export volumes, portfolio diversity etc.) as inputs
- Nature of intended prediction
  - ECA claims gain most attention when exceptional → identification of pattern in prior claims
  - Claims for ECAs are an exemption and not many destination countries exhibit claims
  - For a few destination countries and some ECAs claims happen more regularly


MLT data
separate data by year
2007 to 2017 data
2018 data
parameter optimization: apply all combinations of predefined parameter specifications; repeat 10 times
partition data randomly
training data
validation data
train model
assess model performance
continuous target variables
accuracy
Cohen's κ
R²
best parameters
best model
best parameters
best model
best parameters
best model
test model performance against 2018 data

MODEL BENCHMARK AND MODEL PERFORMANCE ASSESSMENT

- ML approaches are complex and resource-intensive but might not achieve significantly better results (England and Verrall 2002)
- Therefore, we evaluate their performance against a simple benchmark (BM) technique, based on claims ratio of an ECA

$$\hat{r}_{i,j,t} = \frac{\sum_{v=1}^{l} c_{i,j,t-v}}{\sum_{v=1}^{l} e_{i,j,t-v}}.$$

- Transformation of BM estimator into a binary YES/NO variable or a claim ratio class
- Model Performance assessment:
  - Accuracy of prediction to test data
  - Cohen's κ

$$\kappa = \frac{p_o - p_c}{1 - p_c},$$

AGENDA

01 Introduction and Motivation

02 Background on Export Credit Insurance and Claims Prediction

03 Introduction to Machine Learning Approaches and Dataset

04 Results and Discussion

05 Conclusion and Outlook

RESULTS

Table 4. Best parameter and best model results: Accuracy (bold: best performing ML technique).

Task	Outcome	Dataset	DT	RF	NN	PNN	*BM*
Claims YES/NO	Best parameters	Validation	0.886	**0.900**	0.887	0.881	*0.901*
		Test	0.878	0.889	0.874	**0.897**	*0.896*
	Best model	Validation	0.900	**0.909**	0.900	0.898	*0.901*
		Test	0.878	**0.890**	0.848	0.864	*0.896*
Claim ratio class	Best parameters	Validation	0.881	**0.888**	–	0.877	*0.867*
		Test	0.861	0.869	–	**0.888**	*0.858*
	Best model	Validation	0.896	**0.903**	–	0.897	*0.867*
		Test	0.864	**0.870**	–	0.855	*0.858*

Table 5. Best parameter and best model results: Cohen's κ (bold: best performing ML technique).

Task	Outcome	Dataset	DT	RF	NN	PNN	*BM*
Claims YES/NO	Best parameters	Validation	0.352	**0.439**	0.357	0.292	*0.566*
		Test	0.322	**0.408**	0.340	0.275	*0.578*
	Best model	Validation	0.421	**0.489**	0.433	0.358	*0.566*
		Test	0.297	**0.423**	0.303	0.284	*0.578*
Claim ratio class	Best parameters	Validation	0.252	**0.336**	–	0.211	*0.446*
		Test	0.250	**0.320**	–	0.175	*0.458*
	Best model	Validation	0.276	**0.392**	–	0.272	*0.446*
		Test	0.240	**0.336**	–	0.170	*0.458*

Table 6. Best parameters and best model results: R^2 (bold figures: best performing ML technique).

Task	Outcome	Dataset	DT	RF	NN	PNN	*BM*
Claim ratio	Best parameters	Validation	0.038	**0.071**	0.066	–	*0.000*
		Test	0.021	**0.053**	0.046	–	*0.011*
	Best model	Validation	0.081	**0.128**	0.126	–	*0.000*
		Test	0.037	**0.074**	0.027	–	*0.011*

RESULTS

- Amongst the ML techniques, with only two exceptions RF generate the best performance
  - The accuracy achieved against the “Claim ratio class” task is not much different from the accuracy of the less challenging “Claims YES/NO” task. However, Cohen’s κ is more reflective of performance differences, indicating that both, validation and test performance, deteriorate as the task becomes more difficult.
- None of the investigated ML techniques yield satisfactory results against the “Claim ratio” task
  - predictions of actual claim ratios turned out to be largely unreliable.
- The test performance is lower than validation performance (with only two exceptions), performance losses more pronounced when measured by Cohen’s κ
- No definitive conclusion can be made, whether
  - Should validation identify optimal model parameters
  - or generate the specific model for prediction
    sometimes utilizing the best parameters, sometimes employing the best model yields better test performance
- ML techniques and BM perform similar
  - In terms of Cohen’s κ : BM performs better than any of the ML techniques
    some ECAs experience uninterrupted sequences of claims with certain destinations. Therefore, the simple rule “claims in $t-1$ indicate claims in t” employed by the BM works well against the “Claims YES/NO” task, and also against the “Claim ratio class” task
  - Against the “Claim ratio” task, the ML techniques outperform the BM, although at a very low level

RESULTS

Kruskal-Wallis tests on ML technique performance (test data; bold figures: highest median rank)

			Median Rank			
Task	**Measure**	***p*-Value**	**DT**	**RF**	**NN**	**PNN**
Claims Y/N	Accuracy	0.0	11,628.5	**19,613.5**	8972	10,504
	Cohens κ	0.0	11,134.5	**19,716.5**	13,003	5144
Claim ratio class	Accuracy	0.0	8365	**11,558.5**	–	4934
	Cohens κ	0.0	6526.5	**11,467.5**	–	4394
Claim ratio	R^2	0.0	2425.5	**11,310.5**	6115.5	–

RESULTS

Correlation and relationship between validation and test performance

Task	Measure	ML Technique	Validation-Test Correlation	Intercept (Std. Error)	Slope (Std. Error)
Claims Y/N	Accuracy	DT	0.981	−0.078 (0.003)	1.073 (0.004)
		RF	0.990	−0.097 (0.002)	1.098 (0.003)
		NN	0.952	−0.102 (0.003)	1.079 (0.004)
		PNN	0.990	−0.319 (0.002)	1.346 (0.002)
	Cohen's κ	DT	0.851	0.045 (0.002)	0.882 (0.009)
		RF	0.905	0.020 (0.003)	0.970 (0.008)
		NN	0.492	0.141 (0.003)	0.504 (0.010)
		PNN	0.688	0.090 (0.002)	0.625 (0.008)
Claim ratio class	Accuracy	DT	0.976	−0.108 (0.004)	1.107 (0.004)
		RF	0.979	−0.154 (0.004)	1.159 (0.004)
		PNN	0.978	−0.320 (0.003)	1.346 (0.004)
	Cohen's κ	DT	0.902	0.025 (0.001)	0.882 (0.007)
		RF	0.908	0.017 (0.002)	0.924 (0.007)
		PNN	0.882	0.017 (0.001)	0.830 (0.006)
Claim ratio	R^2	DT	0.214	0.011 (0.000)	0.168 (0.017)
		RF	0.706	0.013 (0.001)	0.812 (0.018)
		NN	0.611	0.007 (0.000)	0.487 (0.007)

AGENDA

01 Introduction and Motivation

02 Background on Export Credit Insurance and Claims Prediction

03 Introduction to Machine Learning Approaches and Dataset

04 Results and Discussion

05 Conclusion and Outlook

CONCLUSION AND OUTLOOK

- Comparison of 4 ML techniques
  - Random Forest provide best results across a range of measure
  - Poor performance of several ML for most challenging prediction task “claims ratio” is not surprising, but the large difference in quality of performance compared to the other two tasks is

    Potentially due to features of claims data, such as (very) low frequency/ high severity claims

- “traditional” econometric methods help to extract relationships from masses of data and reveal interdependencies between variables
  - Most ML techniques, incl. RF, NN and PNN are black boxes (Olden and Jackson 2002)

- What is the contribution of techniques that help to better predict but not to better understand a subject (in academia and practice)
  - Decision Trees are an exception

- → we recommend to use RF supported by DT alongside

CONCLUSION AND OUTLOOK

BUT

- The ML models do often not perform better (or even worse) than the simple Benchmark method
  - If ECA transaction to a country had already been insured in the past, the BM was the best predictor
  - If ECA did cover a destination country for the first time, ML could be used as an alternative then they did predict as well as if there was experience with the destination country

Outlook

- Comparison of ML techniques to traditional insurance claims prediction methods such as Chain-Ladder or Bornhuetter-Ferguson

- Integrating problem specific models with ML techniques, e.g. probabilistic distributions for low-default portfolios similar to our the characteristics of our dataset

THANK YOU

Looking forward to your questions, comments and recommendations

simone.krummaker@city.ac.uk