

---

# Actuaries and Data Scientists in the Insurance Industry

Küsnacht, 19. Juni 2018

Junge Aktuare SAV

# Inhaltsverzeichnis

- (1) Einleitung: Umfeld
- (2) Data Science: Definitionen und Vergleiche
- (3) Implikationen für Aktuare
- (4) Tutorial(s) der Fachgruppe
- (5) Beispiele (Swiss Re)
- (6) Fachgruppe Data Science



# Einleitung

Patrizia Kern, Head Marine, Swiss Re Corporate Solutions<sup>(1)</sup>

## Neue Probleme ahoi

Die Vernetzung  
und Beschleunigung  
im Bereich Marine  
nimmt rasant zu.  
Fortschritte in der  
Technologie steigern  
nicht nur die Effizienz,  
sondern bergen auch  
neue Gefahren.

VON PATRIZIA KERN

Die Transportindustrie steht vor einem tiefgreifenden Wandel. In 15 Jahren werden wir sie kaum wiedererkennen. Zum einen werden die Handelsvolumina signifikant steigen, bis 2030 ist beinahe eine Verdreifachung des Handelsverkehrs zu erwarten. Haupttreiber hierfür sind die demografische Entwicklung in Süd-Asien, der Aufstieg der Mittelschichten in Indien, Indonesien und China sowie die Entstehung von modernen Grossstädten (Smart Cities), welche die globale Nachfrage nach Rohstoffen und den Güterhandel exponentiell ansteigen lassen werden.

Zum anderen wird der technologische Fortschritt eine nahtlose Vernetzung der globalen Marine-Lieferkette ermöglichen, wenn nicht sogar erzwingen. Dafür verantwortlich ist die rasante Entwicklung im Bereich der Sensorik, welche gekoppelt mit der Miniaturisierung und Integration neuer Übertragungstechniken eine permanente digitale Kommunikation zwischen der Herstellung von Gütern, deren Handel, Bereitstellung und Nutzung ermöglichen wird.

Der Weg zu intelligenten Schiffen (Smartships) sowie zu (teil-)autonomen Strassen- und Wasserfahrzeugen ist nicht mehr weit. Diese neue Art von Fahrzeugen, gekoppelt mit automatisierten Beladungssystemen, wird nicht nur die Wertschöpfungskette der Schifffahrt erheblich verändern, sondern auch die entsprechende gesamte Risikolandkarte. Es ist davon auszugehen, dass die Seefahrt durch Effizienzsteigerungen sicherer wird, jedoch aufgrund gesteigerter Kapazitäten, Verdichtung des Schiffsverkehrs und Mengen an transportierten Waren die Transportrisiken bedeutend zunehmen werden. Hingegen werden wiederkehrende kleinere

Risiken aufgrund der Möglichkeit von vorausschauender Instandhaltung (Predictive Maintenance) sowie rechtzeitiger präventiver und Ad-hoc-Vorbeugungsmaßnahmen reduziert werden können.

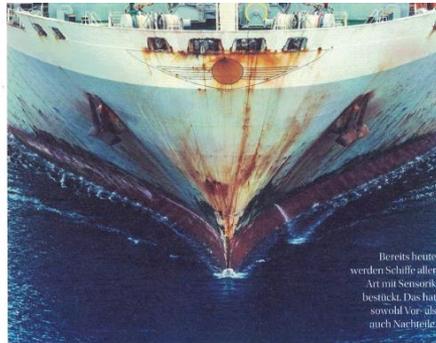
Gerade die Ermöglichung von genaueren Risiko-Szenarien mittels Big-Data- & Smart-Analytics-Techniken stellt eine besondere Herausforderung dar. Die Anzahl an zu untersuchenden Kausalzusammenhängen, der Zugang, die Sammlung, Bereinigung und Analyse von geeigneten Daten sowie die Gewinnung von signifikanten Merkmalen für die Entwicklung von neuen, vorausschauenden Risiko-Modellen (Predictive Modelling) erfordern die Vernetzung und das Zusammenspiel einer Vielzahl unterschiedlicher Engineering-, Analytics- und Technologie-Disziplinen.

**Digitale Transformation**  
Die Vernetzung und Beschleunigung der Transportindustrie nimmt also rasant zu. In diesem neuen digitalen Kontext werden Ver-

sicherer vor einem gewaltigen transformativen Veränderungsprozess stehen, da ihre Wertschöpfungskette stark herausgefordert, wenn nicht sogar zertrennt wird. Wieso ist dem so? Zunächst müssen wir verstehen, dass die Leistungserstellung der Versicherung vereinfacht ausgedrückt auf der Erfassung von Risiko-Informationen, der Modellierung der Risiko-Szenarien, der Ermittlung der erwarteten Schadenhöhe und der Schadenfrequenz zur Berechnung der jeweiligen Risikoprämie basiert. Traditionell stützen sich Versicherer primär auf verhältnismässig wenig Informationen über den zu versicherten Gegenstand. Anhand historischer Schadenereignisse wird ein statistisches Risiko abgeschätzt.

In digitalen Zeitalter werden Sensor-Daten eine dynamische Betrachtung des Risikos – zum Beispiel durch eine kontinuierliche Überwachung der Zustände der bewegten Güter gekreuzt mit Navigationsdaten, Verkehrsdaten und Wetterdaten – in Echtzeit ermöglichen. Diese neue Ausgangslage wird nicht nur die Leistungserstellung der Versicherung verändern, sondern auch die Leistungserbringung.

Die sich entwickelnde digitale Lieferkette, die primär auf dem Sammeln und Auswerten von Echtzeitdaten basiert, wird die Teilnehmerlandkarte und die Machtverhältnisse vollkommen verändern. Ein Beispiel: Wenn Risiko-Modellen (Predictive Modelling) erfordern die Vernetzung und das Zusammenspiel einer Vielzahl unterschiedlicher Engineering-, Analytics- und Technologie-Disziplinen. haben, werden morgigen Reederei, Schiffbetreiber, Logistikler sowie Industriehersteller ihre Betriebsrisiken besser einschätzen können. Dadurch werden sich die Versicherungs-nachfrage und der Deckungsbedarf signifikant verändern.



Bereits heute  
werden Schiffe aller  
Art mit Sensorik  
bestückt. Das hat  
sowohl Vor- als  
auch Nachteile.

In diesem Kontext müssen Versicherer ihre Kernkompetenzen im Bereich der Analyse von gewaltigen Datenströmen zur Gewinnung neuer Erkenntnisse und Einblicke in der sich laufend verändernden globalen Risikolandkarte fundamental ausweiten. Zudem müssen sie zur Sicherstellung ihrer Beteiligung und Mitwirkung Kooperationen mit Dateneigentümern, Datenanbietern sowie Technologieanbietern eingehen und ausweiten, um Wettbewerbsvorteile und – unter Umständen – Existenzen sicherzustellen.

### Konkurrenz von ausserhalb

Digitale Transformationen, ungeachtet der Branche, öffnen eine Vielzahl neuer Opportunitäten und Geschäftsmodelle auch für branchenfremde Akteure. Das gilt vor allem für jene, die eine breite Kundenreichweite, Kundennähe und Kundenbindung bzw. einen weitgreifenden technologischen Vorsprung besitzen. Ausschlaggebend wird künftig jedoch nicht nur das Beherrschen von

neuen technischen Fähigkeiten sein, sondern auch die Bereitschaft, in einem kapitalintensiven Geschäftsbereich einsteigen zu wollen und zu können.

Das Vorhandensein von Daten sowie die Möglichkeit und Fähigkeit, diese in Erkenntnisse (Insights) umzuwandeln, werden sich rasch ausbreiten und diese Nutzung wird sich entlang der Kernkompetenzen eines Unternehmers entfalten. So wird die Fähigkeit, strategische Partnerschaften mit Kunden, Daten-Providern und Technologie-Partnern eingehen zu können, zunehmend auch für Versicherer ausschlaggebend sein, um das digitale Eco-System zu erschliessen. Durch kontinuierliche Innovationen kann so ein Vorsprung geschaffen bzw. profitabel gewachsen werden.

Patrizia Kern ist Head Marine bei Swiss Re Corporate Solutions.

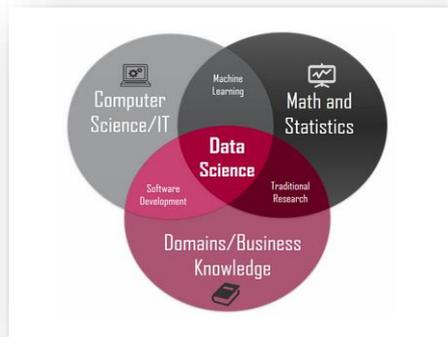
### Kernaussagen:

- «Traditionell stützen sich Versicherer primär auf verhältnismässig wenig Informationen über den zu versicherten Gegenstand.»
- «[...] Zusammenspiel einer Vielzahl unterschiedlicher Engineering-, Analytics und Technologie-Disziplinen.»
- « [...] Sammeln und Auswerten von Echtzeitdaten [...] »
- « [...] ihre Kernkompetenzen im Bereich der Analyse von gewaltigen Datenströmen zur Gewinnung neuer Erkenntnisse und Einblicke in [...] Risikolandkarte [...] »
- «[...] Beherrschen von neuen technischen Fähigkeiten sein [...] »

<sup>(1)</sup> Risk Management Guide 2018, 02.05.2018

# Data Science

Definition(en) und Unterschiede Data Science / Actuarial Science<sup>(2)</sup>



... Schnittmenge von ...

- Mathematik und Statistik
- Computerwissenschaften / Informatik
- Experten- / Fachwissen

... oder Schnittmenge von ...

- Statistischer Modellierung bzw. Maschinellem Lernen
- Hoher Rechenleistung
- Verwaltung und Speicherung von Massendaten
- Experten-/Fachwissen

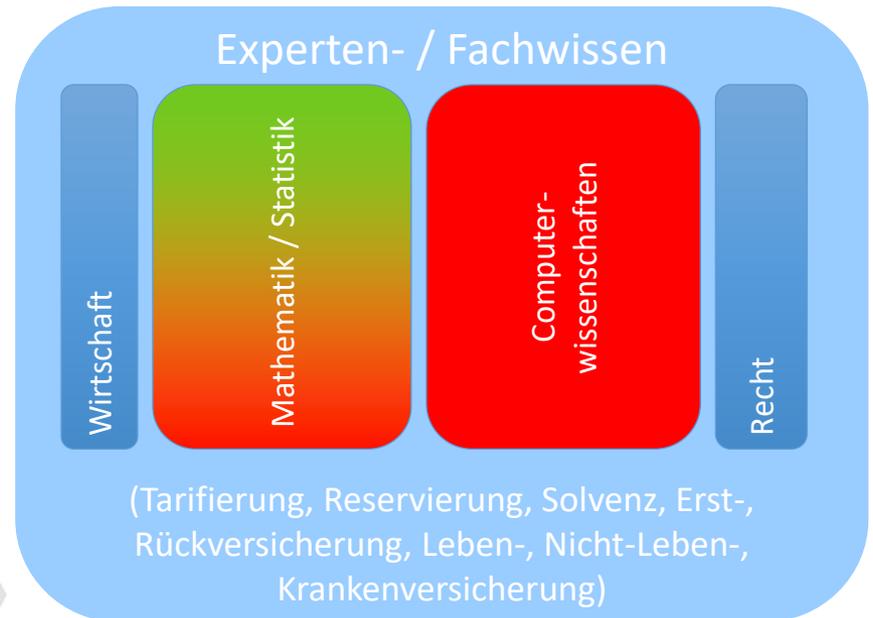
	Aktuar-Wissenschaften	Daten-Wissenschaften
<b>Grundlagen</b>	Mathematische Grundlagen	
<b>Daten</b>	Kleindaten	Klein- und Massendaten
	Strukturierte & statische Daten	Unstrukturierte & dynamische Daten
	Interne Daten	Externe Daten
<b>Mathematik &amp; Statistik</b>	Wahrscheinlichkeitstheorie	Rechnergestützte Statistik
	Leben- und Nicht-Leben Vers.-Math.	Algorithmen
	Quantitatives Risiko Management	Informationstheorie
<b>Computerwissenschaften</b>		Maschinelles Lernen & Visualisierung
		Numerische Optimierung
		Datenmanagement
<b>Programmiersprachen</b>	SAS, S Plus, R	Python, R
	SQL	SQL
	Excel / VBA	Julia, Spark, Scala
<b>Fachwissen</b>	Reservierung, Tarifierung	
	ERM, ALM, Solvenz	
	Buchhaltung, Ökonomie, Recht	

<sup>(2)</sup> Aus: Strategie im Bereich Data Science, Fachgruppe Data Science, SAV, Version V0.9, Mai 2018

# Implikationen für Aktuare<sup>(2)</sup>

Die externen Veränderungen implizieren folgendes:

- Beibehaltung des Fach- und Expertenwissen.
- Fundamentale Änderungen der verwendeten Skills und Fähigkeiten.
- Enge Zusammenarbeit in Kernbereichen der Aktuare mit Data Scientist und Software Engineers.



<sup>(2)</sup> Aus: Strategie im Bereich Data Science, Fachgruppe Data Science, SAV, Version V0.9, Mai 2018

# Tutorial(s) der Fachgruppe

Schadenfrequenz: Vergleich GLM, Regression Tree, Boosting and Neural Network<sup>(3)</sup>

- **Daten:**

- French Motor Third Party Liability Portfolio ( R Datensatz freMTPL2freq)
- Ca. 680'000 Policen
- Anzahl Schäden pro Police
- Exposure, Area, VehPower, VehAge, DriveAge, BonusMalus, VehBrand, VehGas, Region

- **Verglichene Modelle:**

- GLM, Regression Tree, Boosting Machine, Neural Network

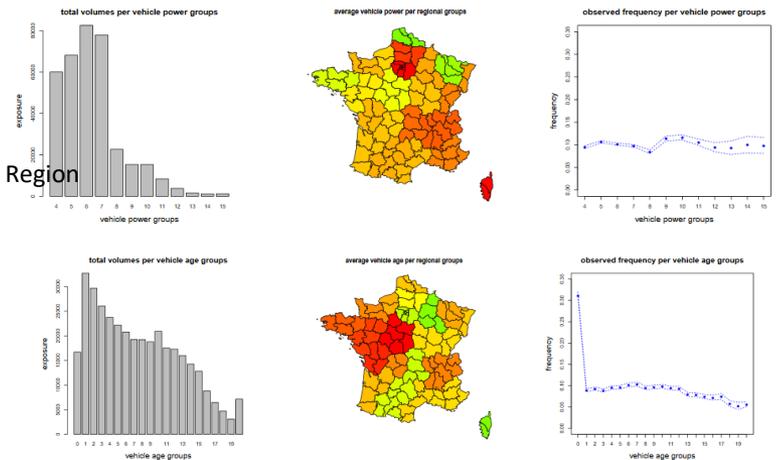
- **Schlussfolgerungen:**

- GLM berücksichtigen keine Interaktionen zwischen den Variablen, während die anderen Modelle diese modellieren.
- Neural Network, Boosting Machine and Regression Tree haben einen kleineren Vorhersagefehler

- **Bemerkungen:**

- Einfach verständliches Tutorial für alle Aktuare.
- Öffentlich zugängliche Daten.
- R Code und Tensorflow-Implementation werden zur Verfügung gestellt
- *Weitere Tutorial werden folgen:*

[https://www.actuaries.ch/de/ueber-uns/fach-arbeitsgruppen/fachgruppe\\_data-science/aktuelle-aktivitaeten](https://www.actuaries.ch/de/ueber-uns/fach-arbeitsgruppen/fachgruppe_data-science/aktuelle-aktivitaeten)



	in-sample loss	out-of-sample loss
Model NN (shallow with $q = 20$ )	30.45048	31.58770
Model PBM3 (depth $J = 3$ , iterations $M = 50$ )	30.09093	31.41314
Model RT2 (min. CV rule, 33 leaves)	30.70841	31.69326
Model GLM1	31.26738	32.17123

Table 10: in-sample and out-of-samples losses of the neural network predictions, the boosting machine, the regression tree and of the GLM approach, units are in  $10^{-2}$ .

<sup>(3)</sup> Case Study: French Motor Third-Party Liability Claims, A. Noll, R. Salzmann, M. V. Wüthrich, 2018, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3164764](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3164764)

---

# Fragen und Diskussion