

g-COM Insurance

galor – Cluster Optimization Model Insurance:
The solution for improved risk analysis and tariffing

Challenge

The insurance or credit risk of an individual customer typically depends on the customer's individual characteristics. The goal is to provide an accurate quantitative classification and sound reasoning for the prediction of future risk.

Benefits

The main benefit is improved prognosis of insurance or credit risk leading to a better tariffing.

The quality and accuracy of the analysis can be employed to yield improved revenues, broader market penetration, and better marketing position.

Data mining and analytics techniques have become main stream not only in the analysis of credit and insurance. The predominant task in the analysis of massive data is the identification of (statistically) relevant structures and a subsequent valid prediction of future developments.

Extracting knowledge from insurance data

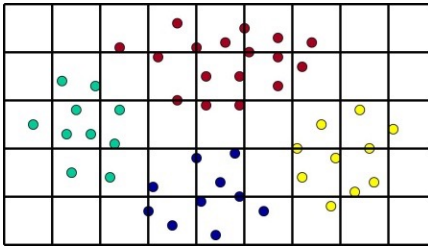
Analyzing insurance data with the highest level of accuracy is important for an optimal prognosis of risk. With **g-COM Insurance**, the customization of **g-COM** for insurance risk predictions, the risk can be analyzed for different groups of similar customers. In contrast to existing approaches, the characteristics of these groups will not have to be predefined but are determined by a machine learning algorithm.

Actuaries can now analyze the groups and their characteristics and use their expertise for explaining the behavior of each group. Based on this detailed analysis the future risk of an existing or new customer can be predicted with much higher precision than before.

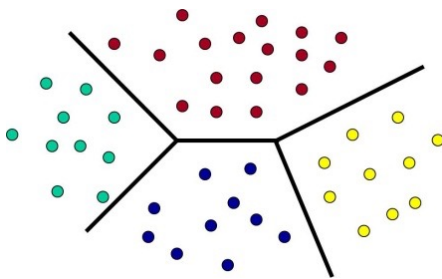
As an example, in a recent study in the realm of residential building insurance, the homeowners risks related to water pipe leaks had to be classified. Using **g-COM Insurance** relevant features of the data could be identified. For instance, smaller homes built less than 30 years ago, in the storm and water zone 2 bore the same risk as up to 60 years old larger houses in storm and water zone 1. The risk of smaller homes built up to 60 years ago in storm zone 2 and water zone 3 was, however, 80% higher.

Disadvantages of standard methods

Standard approaches typically employ pre-segmentation techniques that are based on the independent analysis of the influence of all relevant parameters. Geometrically, this leads to a dissection of the parameter space into boxes



Sparsely populated or empty cells



Significant substructures

which contain more homogenous substructures that are then subject to statistical analysis. In the presence of a reasonably large number of parameter characteristics and specifications the generated cells are typically sparsely populated. Hence, quite often, an application of the law of large numbers is prohibited and different, more complicated and less reliable, statistical techniques have to be evoked.

g-COM: A change in paradigm

g-COM uses a reverse, data driven paradigm. Rather than imposing a simple dissection on the parameter space followed by complicated statistics **g-COM** computes an optimal clustering of this space and subsequently applies simple, meaningful and reliable statistics. Until recently such an effective and efficient application of this natural principle was out of reach due to the lack of an adequate mathematical model and fast algorithms for clustering of high dimensional weighted data, including nominal data, under all relevant problem specific constraints. While standard clustering methods are capable of determining rather homogenous structures efficiently, such additional constraints could not be incorporated appropriately. Further, standard methods are typically restricted to computing solutions that are merely locally optimal, while the new technique is capable of creating globally optimal solutions. **g-COM** therefore allows to gain new insight in the dependencies and interactions of all parameters for the respective application by allowing for a significant analysis of the clusters at optimality.

g-COM in a nutshell

g-COM, the galor Cluster Optimization Model, first solves an application and data specific complex clustering problem with the aid of state-of-the-art and newly developed mathematics and software technology. **g-COM** thereby identifies homogenous substructures within the total body of data and performs a robust and statistically significant analysis of the multivariate interactions of all parameters. This is the bases of the **g-COM** prediction method which can then be used for the classification of new data.

Knowledge discovery with g-COM Insurance

With **g-COM Insurance** the typically inhomogenous population of insurance customers is split into a small

number of groups which behave very similarly with respect to their inflicted risk. In contrast to classical sub-group analysis this multivariate approach guarantees the detection of large enough but still quite homogeneous populations so as to warrant significant results that can be assessed by the actuary's favorite tool. **g-COM Insurance** is able to determine representatives of the relevant groups and provide in depth analysis. Combined with the knowledge and experience of actuaries this yields a new and precise way of interpreting and predicting risk. **g-COM Insurance** comes with various front end tools including a competitive scenario analyzer that allows to simulate the effect of a **g-COM** based tariffing under realistic market assumptions set by the user. Figure 1 shows a competitive scenario for a car insurance application.

	classical tariffing	G-COM
Price level (%)	100,00	100,00
Number of customers	1.400.739	1.517.736
Total revenue (€)	337.159.503	435.635.532
Total costs (€)	352.698.705	421.159.843
Total profit (€)	-15.539.201	14.475.689
Return (%)	-4,41	3,44
Revenue per customer (€)	240,70	287,03
Costs per customer (€)	251,79	277,49
Profit per customer (€)	-11,09	9,54

Figure 1: Competitive Insurance Scenario

g-COM Insurance within the realm of business analytics

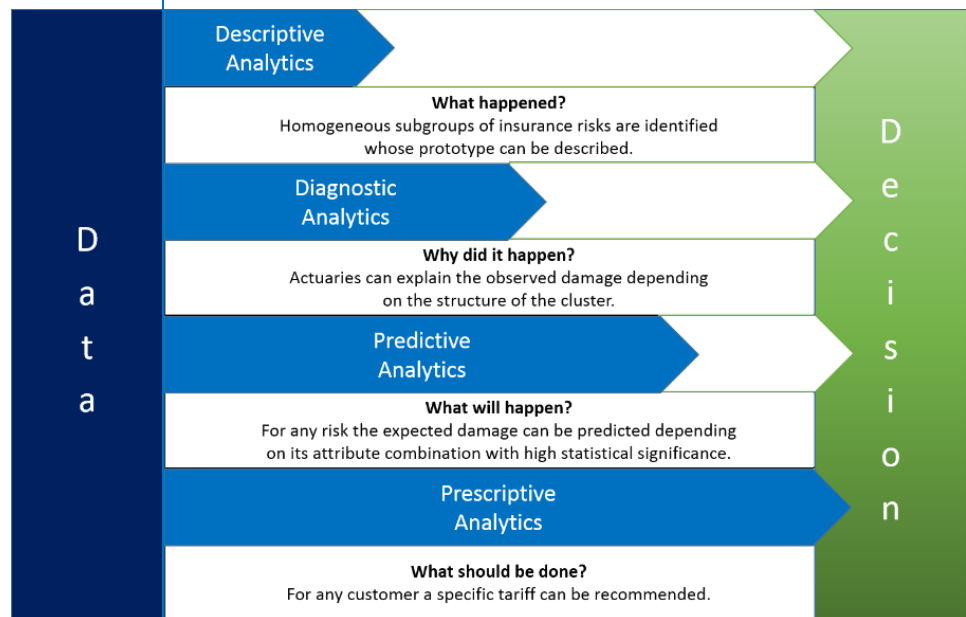
g-COM Insurance can be described within the realm of the phase model of business analytics:

What happened? This task is performed by identifying and analyzing homogeneous substructures and their prototypes. **g-COM Insurance** reveals these structures in a form that is directly accessible by actuaries or risk managers.

Why did it happen? On the basis of the structural information risk experts gain new insight in the correlation between multiple attributes and the resulting risk, potentially leading to new explanations.

What will happen? Based on the identified structures within the high-dimensional space of attributes **g-COM Insurance** provides significant predictions for the risk inflicted upon the insurance company by any individual customer. In fact, his attribute combination places a new

customer into a reference group whose past extent of damage is known and whose future risk can be determined with high accuracy.



Inspired by and in style of Gartner, #G00254653 (September 2013)

What should be done? **g-COM Insurance** can be used to derive a new tariffing system. But it can also be applied in a "noninvasive" form i.e., without the need of major changes in business operations, by producing merely correction factors for the existing tariff.

Unlike other methods, **g-COM Insurance** does not function as a black box but allows for reasoning.

Are you ready to start a project?

To assess the potential of **g-COM** in a new field of application the following is needed:

- A short description of the application including the desired analytic goals;
- A list of the relevant parameters and their specifications;
- A representative set of test data.

Contact us!

Gesellschaft für Angewandte Logik und Operations Research
 Spitzelbergstraße 10a, D-81476 Munich, Germany
 Managing Directors: Dr. Peter Gritzmann, Dr. Andreas Brieden
 Phone ++49 / 89 / 74 54 53 70 • Fax ++49 / 89 / 74 54 53 80 • www.galor.de • g-com@galor.de

General applicability

g-COM can be applied to a great variety of data analysis tasks in various business sectors. It has already been proven highly successful in the prediction of credit and insurance loss, air cargo demands, and medical treatment efficacy.

