Fachhochschule Dortmund
University of Applied Sciences and Arts

Fachbereich Wirtschaft

Forschungsstelle
Risikomanagement, Versicherungsregulierung und Vertrieb
Prof. Dr. Michael Radtke ● Prof. Dr. Matthias Beenken

Mitteilung 2/2019

2. Data Science Challenge

Veranstaltung vom 4./5.12.2019 an der Fachhochschule Dortmund

Prof. Dr. Matthias Beenken

Jessica Michalczyk

Prof. Dr. Jens Mörchel

1. Idea

Our world is increasingly data driven. Data is generated and stored in unimaginable quantities, sometimes structured, but increasingly also unstructured. Mobile phones, fitness trackers, and motor vehicles, just to name three examples, are constantly producing large amounts of unstructured data (Big Data), which can be very interesting for many companies. They help to better understand the behavior and needs of consumers.

Insurance companies increasingly need to understand and make sense of this data. This would enable them to know how their customers behave and what risks they are exposed to. As a result, it is possible to take more data driven business decisions, e.g. the setting of insurance rates according to individual driving habits.

Often, however, things are different. Many companies, including insurance companies, have a wealth of proprietary, structured data that they barely use. By using this data they would know a lot more about their customers. Personal data, contract data and claims data represent a treasure of information about the customer, which has not been raised by many insurers yet. Next to competitive reasons, such increased knowledge of markets and customers is also required by the regulatory environment.

In order to analyze even structured data, it is absolutely mandatory to use on a regular base intelligent statistical tools with a high degree of automation. Therefore insurance companies increasingly need data scientists - people who are able to handle software as well as understand the business model. Data scientists are able to analyze efficiently large amounts of data, which can be used for tariff and product development, pricing or also for sales management etc.

The purpose of the Data Science Challenge is to give students an idea of the importance of Big Data and the development of business models based on data. In addition, students get to know software tools that can be used to efficiently analyze and structure large amounts of data. In addition to the software tools themselves, our partner **IBM** also provides high-caliber trainers who help students with very heterogeneous skills to use these tools in a short time.

While flight data and weather data were used in the first Data Science Challenge 2018, the Dortmund-based insurance company **VOLKSWOHL BUND** provided data from car insurance for the second Data Science Challenge 2019.

Another aspect of the Data Science Challenge is to bring students from different disciplines together and form productive, interdisciplinary teams whose skills complement each other. The participating students learned to master complex tasks in interdisciplinary work environments and use different skills and experiences for their later working lives – programming skills on

the one hand and business skills on the other. Finally, the Challenge had an international background in that both domestic and foreign students were mixed. Therefore, the challenge language was English.

2. Procedure

Participants of the two-day event were over 50 students from the following study programs:

- Bachelor Versicherungswirtschaft dual (Insurance Business)
- Master Financial Management
- European Master in Project Management (international students)
- Master Digital Transformation.

At the beginning of the event, the students were familiarized with the purpose and the process of the Data Science Challenge. Students were given an introduction to IBM SPSS Modeler data mining software.

Subsequently, ten groups were formed, which were given the task in competition to analyze a data set of the VOLKSWOHL BUND insurance and to develop a data-based Use Case.

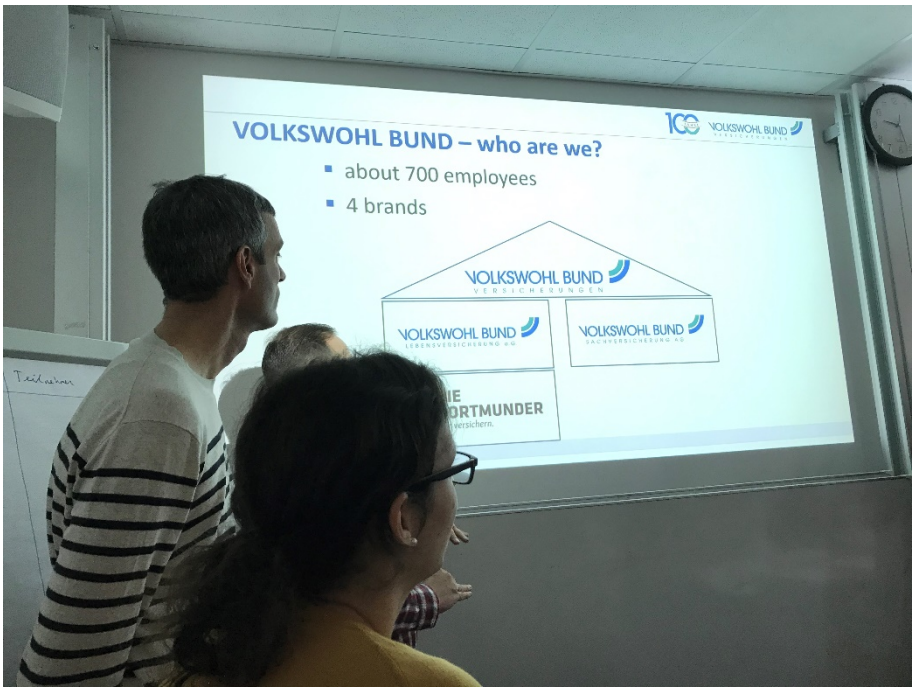

Fig. 1: Initial Presentation of VOLKSWOHL BUND (photo Beenken)

Fig. 2: Initial Presentation of IBM (photo Beenken)

To help students to develop Use Cases, three different possible Use Cases were presented as examples:

1. Calculating tariffs of motor insurance

Initial situation

The insurance company VOLKSWOHL BUND Sachversicherung (VBS) wants to set up a new motor insurance tariff to improve its competitive position. This tariff should target especially attractive customers.

Possible procedure

The dataset contains a large number of tariff-relevant risk characteristics.

The first step is to find out which of these characteristics have any impact on the claims experience and the previous premium calculation. For this purpose, some variables must be recoded, for example, in order to perform correlation and regression calculations. For some variables, discriminant analysis helps to uncover relationships; the simplest form would be cross-tabulations.

Subsequently, data-supported interesting target groups are to be identified, which so far at any rate show sufficient premiums and favorable claims development. It should also be considered which incentives can be given to these target groups, preferably to be insured with the VBS.

2. Product design motor insurance

Initial situation

The insurance company VOLKSWOHL BUND Sachversicherung (VBS) wants to set up a new car insurance tariff with additional services in order to improve its competitive position and retain desirable customers. For this purpose, complementary, for the customer valuable services such as assistance, memberships, payback systems, etc. are to be developed.

Possible procedure

First of all, a deeper understanding of the attractiveness of certain customers and risk relationships must be developed analogously to Use Case 1. The aim is to identify interesting customer groups that are characterized by a low loss ratio and attractive premiums.

Based on this, for example, with the help of a brainstorming or a small design sprint, additional services can be identified that can improve customer loyalty among these customer groups. As far as possible, prices and costs should be determined.

Furthermore, simulations could be used to illustrate how

- Sales of VBS in the motor vehicle sector (booked premiums for motor vehicle total see annual report on the homepage)

- Costs of the VBS (cost ratio of motor vehicle see annual report), taking into account a change in the composition of the stock and

- additional revenue and costs from ancillary services

could develop. For example, a suitable forecast period could be three years. Various parameters, such as cancellation rates, have to be estimated for the simulation.

3. Sales Management of motor insurance: Early Detection System

Initial Situation

As part of its risk management, the insurance company VOLKSWOHL BUND Sachversicherung (VBS) wants to install an early warning system for sales management in order to identify intermediaries who show a noticeable loss trend in their portfolios. Both the frequency of occurrence and the amount of loss can be characterized by coincidences, but may also point to acquisition in differently desirable target groups and relational networks. In rare cases, it is even conceivable that intermediaries at the expense of the company influence the development of claims. The sales management is to be made aware of potential problem cases at an early stage in order to carry out investigations and, if necessary, to be able to review further cooperation.

Possible procedure

The data record contains the broker numbers (hashed). First of all, customers must be identified mediated by each broker and pooled in a portfolio per broker. Next you will find a series of data on claims experience.

First you could analyse which intermediaries have remarkably high and / or conspicuously high damages in relation to their portfolios. In addition, it could be examined whether the

customers within those portfolios have certain characteristics in common that could indicate a specific customer network of the intermediary.

As a result, features or measures could be designed to identify anomalies that could give rise to a review of the broker relationship.

The teams were given the task of examining a data set of portfolios and of claims in motor insurance from a number of years, reorganizing data for research purposes, and exploring relationships between customers, insurance contracts, claims, and brokers. Based on the findings from the data, they should develop their own use case and prove their creativity and knowledge of the application field of motor insurance.

Fig. 3: Working situation (photo Beenken)

The teams were constantly supported by:

- Lena Eckstein, Jochen Stark (IBM)
- Johanna Dziobek, Patrick Nielinger (VOLKSWOHL BUND)
- Prof. Dr. Helena Lovasz-Bukvova (Fachhochschule Krems)
- Prof. Dr. Christian Reimann (Fachhochschule Dortmund, FB 4 Informatik)
- Jessica Michalczyk, Prof. Dr. Jens Mörchel, Prof. Dr. Matthias Beenken (Fachhochschule Dortmund, FB 9 Wirtschaft).

Fig. 4: Working situation (photo Beenken)

Fig. 5: Working situation (photo Beenken)

Fig. 6: Working situation (photo Beenken)

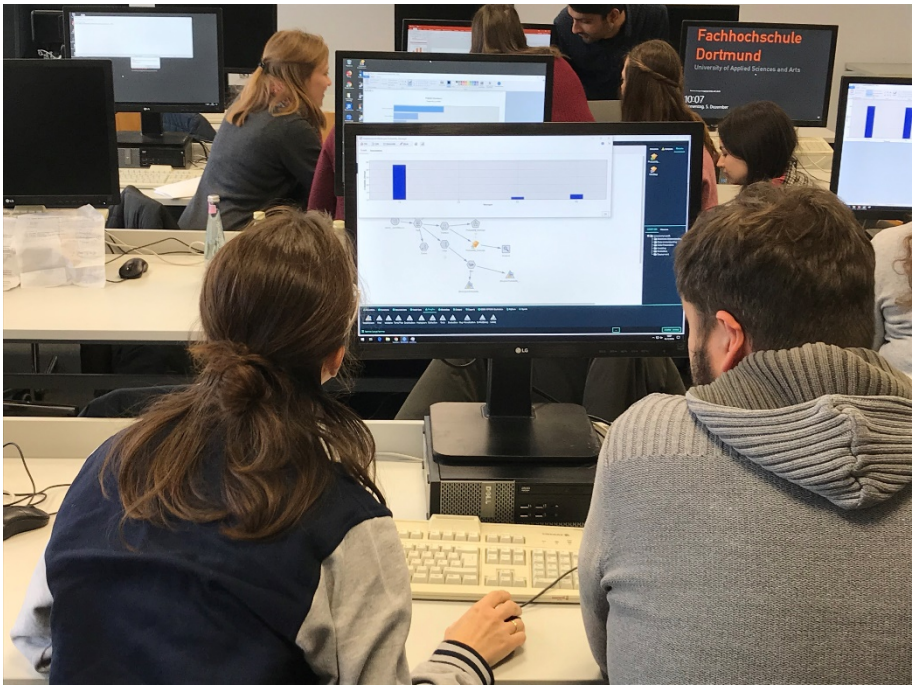

Fig. 7: Working situation (photo Beenken)

On the second day at noon, the teams had to pitch and present their solution. Each team had five minutes to convince the jury of their solution. The guiding questions were:

- What is your Business Case / Solution?
- What did you learn during data analysis?
- What did you learn in general?

Fig. 8: Final instructions (photo Michalczyk)

Fig. 9: Pitch situation (photo Beenken)

Fig. 10: Pitch situation (photo Beenken)

Fig. 11: Jury surveying the pitches (photo Beenken)

Fig. 12: Jury discussing results (photo Beenken)

Fig. 13: Jury explains rating (photo Beenken)

Fig. 14: The winner team does not takes it all (photo Beenken)

Fig. 15: Congratulations to the winner team (photo Michalczyk)

Fig. 16: Farewell (photo Michalczyk)

Fig. 17: Farewell (photo Michalczyk)

4. Results & Learnings

The teams presented very different solutions. Some differences were explained by the fact that the teams had chosen different approaches to clean up the dataset and reorganize the data, and analyzed very different sub-samples. Another factor was the choice of different methods. Predominantly, correlation and regression analyzes were chosen, but some also attempted factor analysis or neural networks.

However, the most important factor was that very different mental models behind the analysis were used. Thus, an important learning effect has been achieved that no intelligent software (e.g. Artificial Intelligence) can replace a deep understanding of the market reality and the business models behind the data.

The groups identified important factors influencing the claims experience in a motor insurance portfolio like

- Customer/Driver: Age, profession, type of housing, region, accompanied driving with 17
- Car: Type, age, usage, garage
- Contract: Type of insurance chosen
- Broker: Portfolio size and claims experience.

Some groups identified interesting target groups that the insurer should focus on, e.g.

- Age-based target groups: Some found younger, some middle-aged groups as most attractive customers.
- Life situation-based target groups: One team proposed to focus on young families who own property and use their car privately. Another team found that business people perform best in motor insurance.
- Car type-based target groups: Several teams found private used passenger cars as most favorable to be insured. One group identified car manufacturers like Peugeot, Opel, VW, and Mercedes to be most attractive. Another team found motorbikes to be a good choice for a motor insurance portfolio.

Some groups developed use cases that focus on retention of favorable customers. In addition to a discount, one team also proposed incentives through goodies and discounts for paperless communication. Another team considered a customer loyalty bonus based primarily on the duration of the contract and estimated the possible amount. A driver training rounded off the incentive design.

Several groups analyzed the brokers´ portfolios and identified some of them being very attractive, others very unattractive. One solution was to offer profit shares to those brokers that deliver particularly good returns.

A lesson learned for all participants was the challenge to work in a mixed team with different academic backgrounds, different cultures and different languages, and to work together to find a result. Some groups solved this through strong leadership by one team leader, others by splitting the work packages into subgroups. The team results were better than what individual participants could have achieved. Anyone who knows the market, but does not master the techniques, will not be able to generate important insights from data. Anyone who masters the techniques but does not know the market and the business models will hardly carry out meaningful analyzes. Only together can we master the digital transformation and develop convincing, digital business models.