Seminar in Business Analytics:
Introduction to Data Mining with R and Python
(Bachelorseminar)
Winter Term 2020/2021

Course Description:

Prior to the start of the Information Age in the late 20th century, companies were forced to collect data from non-automated sources manually. Companies back then lacked the computing capabilities necessary for data to be analyzed, and as a result, decisions primarily originated not from knowledge but from intuition. With the emergence of ubiquitous computing technology, company decisions nowadays rely strongly on computer-aided **"Data Mining"**.

In this seminar, students will learn to build prototypes of such data mining systems targeted at supporting managerial decisions. In more detail, students are typically given a particular data set and first asked to describe and visualize its content of their dataset. Second, students should pick a statistical method / data mining algorithm of their choice and perform a descriptive or predictive data mining task on their dataset.

This is an **introductory seminar targeted at motivated Bachelor students** interested in developing skills in the field of Management of Information systems (Wirtschaftsinformatik), particularly in **programming and data analytics / machine learning**. Bachelor students who have enjoyed and done well in the chair's lectures in Manangement of Information Systems (MIS) and/or Business Intelligence are particularly encouraged to apply. Topics are connected to the chairs research activities, e.g. **Energy Analytics, Urban Analytics, Social Media Analytics, Text Mining and Analytics of financial data**.

In particular, the students will learn to build data mining systems to address the following tasks:

- **Forecasting:** Based on historical values, how can businesses predict future developments ahead of time? Given the current stock market prices, can we predict tomorrow's values?
- **Data analysis:** How does weather impact electricity prices? Which parameters of second-hand cars correlate with their value?
- **Clustering:** How can businesses group consumers into distinct categories according to their purchase behavior? Can businesses group job applicants into groups of similar characteristics?

- **Dimension reduction:** How can businesses simplify a large amount of indicators into a smaller subset with similar significance? Can the huge set of features characterizing supermarkets (e.g. gas station, discounts, service) be combined into groups?

Target Group:

Interested and committed B.Sc. VWL and BWL students. Prior programming skills are welcome, but not a requirement. However, if students do not yet have programming skills, it will be expected from them to familiarize themselves with the necessary skills during the seminar in self-study. Each participant will have to learn to apply a data mining method using the programming languages Python or R.

Organization:

Registration: Until November 5, 2020

Application via email to gunther.gust@is.uni-freiburg.de. **Make sure the following information is contained in the email body**:

- First name, last name
- Matriculation number (Matrikelnummer)
- Current overall grade average
- Attended lectures and seminars at our chair and grade obtained
- Email, phone number
- Study program, semester
- Short description of experience level in Python or R
- Personal preferences regarding one (or several) of potential topics: Energy Analytics, Urban Analytics, Social Media Analytics, Text Mining and Analytics of financial data

In addition, you need to send your **transcript of records in the attachment**.

Response whether application was successful will be sent out shortly after the registration deadline

First meeting:	**Nov 13, 2020 at 16ct (online)**
Paper due:	**Jan 3, 2021**
Presentation:	**Jan 12, 2021, 16 ct.**
Revised paper due:	**March 15, 2021**

Communication:

All announcements, handouts, etc. will be sent via email.

Topics:

Exact topics along with hints on literature will be announced at a later point.

Policies and Procedures

Grading:	Paper (about 15 pages, 50%) and revised final paper (50%). In addition, you have to hand in your programming code and datasets (if applicable). In addition, there will be a presentation session (online), where the participants get feedback on their preliminary work. The seminar paper should be written in English.
Credit points:	6
Credit points are applicable to:	B.Sc. BWL PNPM: Allgemeine BWL B.Sc. VWL: BWL, Wirtschaftsinformatik

Chair:

Prof. Dr. Dirk Neumann
Albert-Ludwigs-Universität
Chair of Information Systems
Platz der Alten Synagoge
79085 Freiburg
Germany