

Advanced Analytics

Mastering Social Media, Web, Satellite and Spatial
Data with R

Prof. Dr. Martin Vogt
Seminar

Big Data und Data Literacy

„Daten sind der Rohstoff der Zukunft.“ (Angela Merkel)

“Big data is at the foundation of all the megatrends that are happening.” (Chris Lynch, Vertica Systems)

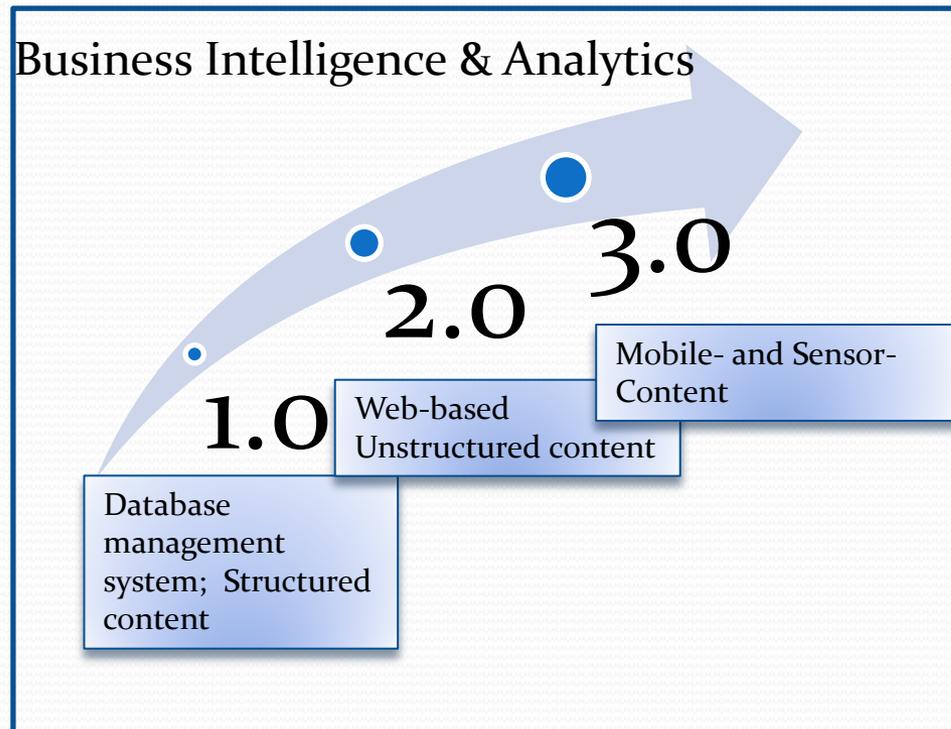
“If we have data, let’s look at data. If all we have are opinions, let’s go with mine.” (Jim Barksdale, former Netscape CEO)

“The ability to take data – to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it’s going to be a hugely important skill in the next decades”. Hal Varian (2009, Chief Economist, Google)



BI und Analytics

- Big Data und Open Data benötigen neue Methoden und Ansätze zur Analyse
- Beispiel: Analyse von Text, Web, Sound, Satelliten oder Spatial Daten
- Daten können in R oder Python eingeladen, analysiert und visualisiert werden
- Bsp. Verbindung zu Twitter mit Paket Rtweet



Seminarstruktur

- Prüfungsleistungen: Je 1 Vortrag (60 Minuten) und 1 Poster oder eine Hausarbeit
- Maximal 20 Teilnehmer
- Individuelles Thema im Zusammenhang mit Datenanalyse. Auf Wünsche wird gerne eingegangen.

Bewertet werden:

- Vortrag: Vortragsstil, Inhalt (eigene Anwendung) und Aufbereitung: 50%
- Poster oder Hausarbeit: DIN A2 : 50% bzw. 10 Seiten
 - m.vogt@wir.hochschule-trier.de
 - Powerpoint oder LaTeX

Seminarstruktur

- Im Folgenden werden nur Beispiele für Themen genannt. Das individuelle Thema wird bilateral besprochen und festgelegt.
- Je nach Wunsch kann das eigene Thema einen Schwerpunkt in einer strategischen Analyse, etwa von Social Media Daten haben (Bsp.: Aufbau einer Social Media Strategie eines fiktiven oder eigenen Unternehmens/Winzers,...).
- Alternativ kann auch ein Verfahren, etwa CNN, LSTM oder Random Forest erarbeitet und angewendet werden.
- Schließlich kann der Fokus auch auf der Programmierung liegen, falls gewünscht.

Beispiel: Textmining (6 Themen)

Introduction to Social Media and Textmining with R

Textmining: Sentiment Analysis with Tidy data

Textmining: Analyzing Word and document Frequency

Textmining: Relationships between words

Textmining: Mining Financial Articles

Textmining: Topic Modeling



Seminarstruktur

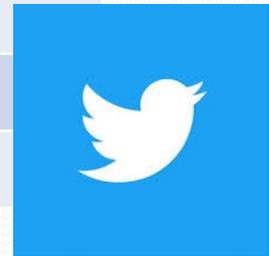
- Im Folgenden werden nur Beispiele für Themen genannt. Das individuelle Thema wird bilateral besprochen und festgelegt.
- Je nach Wunsch kann das eigene Thema einen Schwerpunkt in einer strategischen Analyse, etwa von Social Media Daten haben (Bsp.: Aufbau einer Social Media Strategie eines fiktiven oder eigenen Unternehmens/Winzers,...).
- Alternativ kann auch ein Verfahren, etwa CNN, LSTM oder Random Forest erarbeitet und angewendet werden.
- Schließlich kann der Fokus auch auf der Programmierung liegen, falls gewünscht.

Beispiel: Social and Web Mining (3 Themen)

Mining Opinions, Exploring Trends, and More with Twitter

Scraping Wikipedia and websites

Google Page Rank Algorithm



Seminarstruktur

- Im Folgenden werden nur Beispiele für Themen genannt. Das individuelle Thema wird bilateral besprochen und festgelegt.
- Je nach Wunsch kann das eigene Thema einen Schwerpunkt in einer strategischen Analyse, etwa von Social Media Daten haben (Bsp.: Aufbau einer Social Media Strategie eines fiktiven oder eigenen Unternehmens/Winzers,...).
- Alternativ kann auch ein Verfahren, etwa CNN, LSTM oder Random Forest erarbeitet und angewendet werden.
- Schließlich kann der Fokus auch auf der Programmierung liegen, falls gewünscht.

Beispiel: Spatial Data (5 Themen)

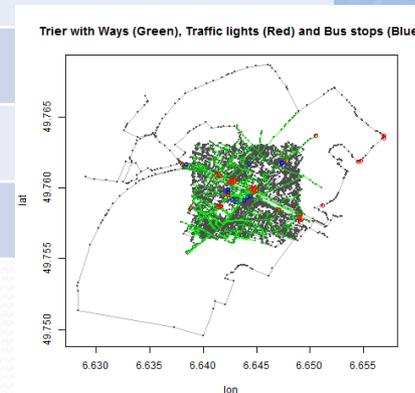
Spatial Data with R

Case Study - How long is the coast of Britain?

Case Study - Distribution of wild potatoes

Spatial analysis with OpenStreetMap

Visualizing Spatial Data



Seminarstruktur

- Im Folgenden werden nur Beispiele für Themen genannt. Das individuelle Thema wird bilateral besprochen und festgelegt.
- Je nach Wunsch kann das eigene Thema einen Schwerpunkt in einer strategischen Analyse, etwa von Social Media Daten haben (Bsp.: Aufbau einer Social Media Strategie eines fiktiven oder eigenen Unternehmens/Winzers,...).
- Alternativ kann auch ein Verfahren, etwa CNN, LSTM oder Random Forest erarbeitet und angewendet werden.
- Schließlich kann der Fokus auch auf der Programmierung liegen, falls gewünscht.

Beispiel: Data Science (5 Themen)

Self-organizing Maps

Support Vector Machine

Neuronale Netze

Random Forrest

Reinforcement Learning

Seminarstruktur

- Im Folgenden werden nur Beispiele für Themen genannt. Das individuelle Thema wird bilateral besprochen und festgelegt.
- Je nach Wunsch kann das eigene Thema einen Schwerpunkt in einer strategischen Analyse, etwa von Social Media Daten haben (Bsp.: Aufbau einer Social Media Strategie eines fiktiven oder eigenen Unternehmens/Winzers,...).
- Alternativ kann auch ein Verfahren, etwa CNN, LSTM oder Random Forest erarbeitet und angewendet werden.
- Schließlich kann der Fokus auch auf der Programmierung liegen, falls gewünscht.

Beispiel: Image Segmentation and Remote Sensing (3 Themen)

Image Segmentation with R – K-Means

Image Segmentation with R – PCA

Satellite Data



Seminarstruktur

- Im Folgenden werden nur Beispiele für Themen genannt. Das individuelle Thema wird bilateral besprochen und festgelegt.
- Je nach Wunsch kann das eigene Thema einen Schwerpunkt in einer strategischen Analyse, etwa von Social Media Daten haben (Bsp.: Aufbau einer Social Media Strategie eines fiktiven oder eigenen Unternehmens/Winzers,...).
- Alternativ kann auch ein Verfahren, etwa CNN, LSTM oder Random Forest erarbeitet und angewendet werden.
- Schließlich kann der Fokus auch auf der Programmierung liegen, falls gewünscht.

Beispiel: Eye Tracking und Sound (2 Themen)

Eye-Tracking einer Webseite (<http://www.eyetracking-r.com/>)

Sound Analysis in R

