

Advanced Analytics

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

Prof. Dr. Martin Vogt
Seminar

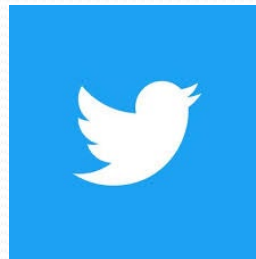
Big Data und Data Literacy

„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)

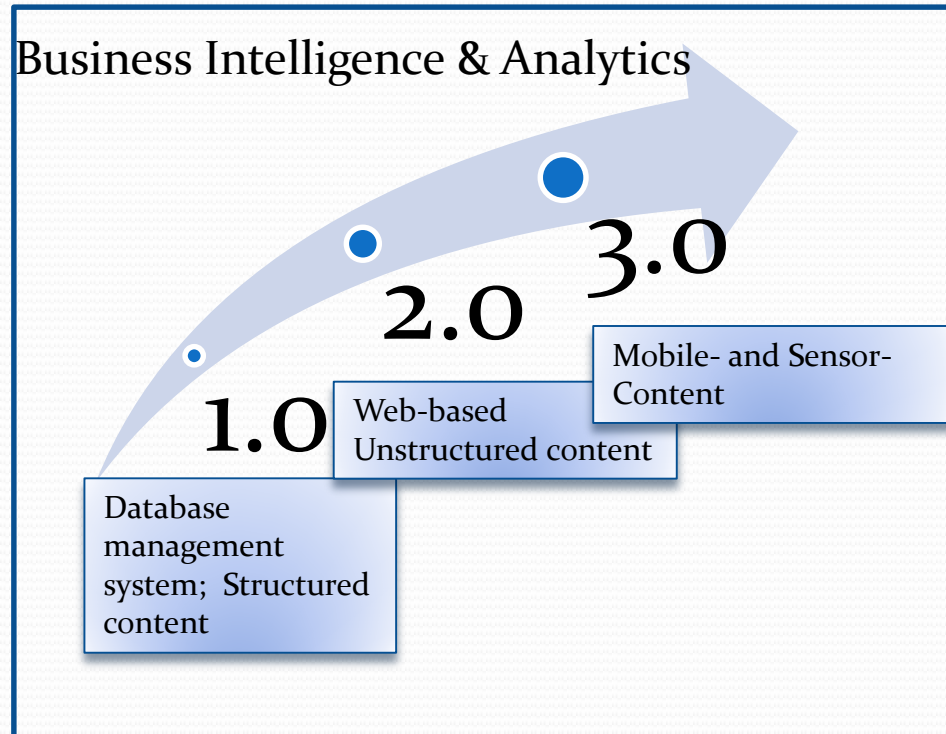
“If we have data, let’s look at data. If all we have are opinions, let’s go with mine.” (Jim Barksdale, former Netscape CEO)

“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, Satelliten oder Spatial Daten
- Daten können in R eingeladen, analysiert und visualisiert werden
- Bsp. Verbindung zu Twitter mit Paket TwitterR oder zu Instagram mit Paket instaR



Seminarstruktur

- Prüfungsleistungen: Je 1 Vortrag (60 Minuten) und 1 Poster
- Maximal 20 Teilnehmer

Bewertet werden:

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

Seminarstruktur

- Prüfungsleistungen: Je 1 Vortrag (inklusive R Umsetzung) und 1 Hausarbeit (circa 10 Seiten)
- Maximal 20 Teilnehmer

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

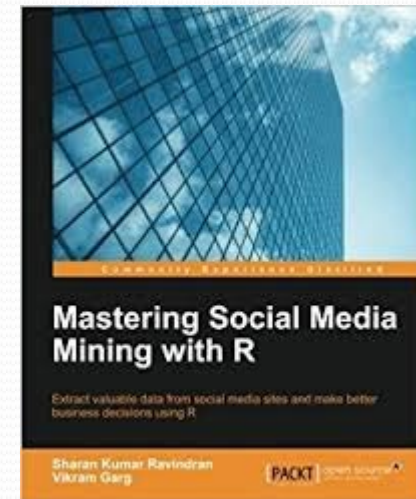
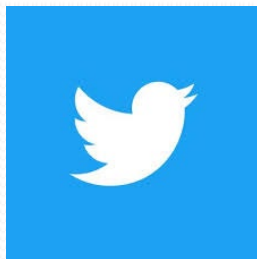
- Prüfungsleistungen: Je 1 Vortrag (inklusive R Umsetzung) und 1 Hausarbeit (circa 10 Seiten)
- Maximal 20 Teilnehmer

Social and Web Mining (3 Themen)

Mining Opinions, Exploring Trends, and More with Twitter

Scraping Wikipedia and websites

Google Page Rank Algorithm



Seminarstruktur

- Prüfungsleistungen: Je 1 Vortrag (inklusive R Umsetzung) und 1 Hausarbeit (circa 10 Seiten)
- Maximal 20 Teilnehmer

Spatial Data (5 Themen)

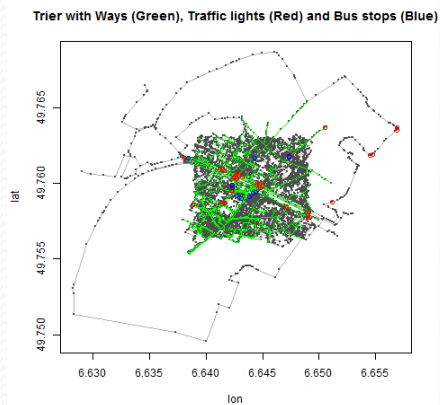
Spatial Data with R

Case Study - How long is the coast of Britain?

Case Study - Distribution of wild potatoes

Spatial analysis with OpenStreetMap

Visualizing Spatial Data



Seminarstruktur

- Prüfungsleistungen: Je 1 Vortrag (inklusive R Umsetzung) und 1 Hausarbeit (circa 10 Seiten)
- Maximal 20 Teilnehmer

Data Science (5 Themen)

Self-organizing Maps

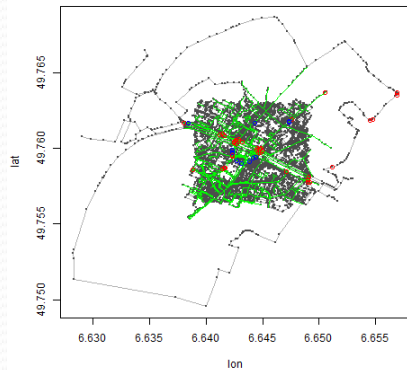
Support Vector Machine

Neuronale Netze

Random Forrest

Reinforcement Learning

Trier with Ways (Green), Traffic lights (Red) and Bus stops (Blue)



Seminarstruktur

- Prüfungsleistungen: Je 1 Vortrag (inklusive R Umsetzung) und 1 Hausarbeit (circa 10 Seiten)
- Maximal 20 Teilnehmer

Image Segmentation and Remote Sensing (3 Themen)

Image Segmentation with R – K-Means

Image Segmentation with R – PCA

Satellite Data



Seminarstruktur

- Prüfungsleistungen: Je 1 Vortrag (inklusive R Umsetzung) und 1 Hausarbeit (circa 10 Seiten)
- Maximal 20 Teilnehmer

Eye Tracking und Sound (5 Themen)

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

Sound Analysis in R

