

Business Analytics

0. Overview

Lars Schmidt-Thieme

Information Systems and Machine Learning Lab (ISMLL)
University of Hildesheim, Germany

Outline

1. What is Business Analytics?
2. Overview of the Lecture
3. Organizational Stuff

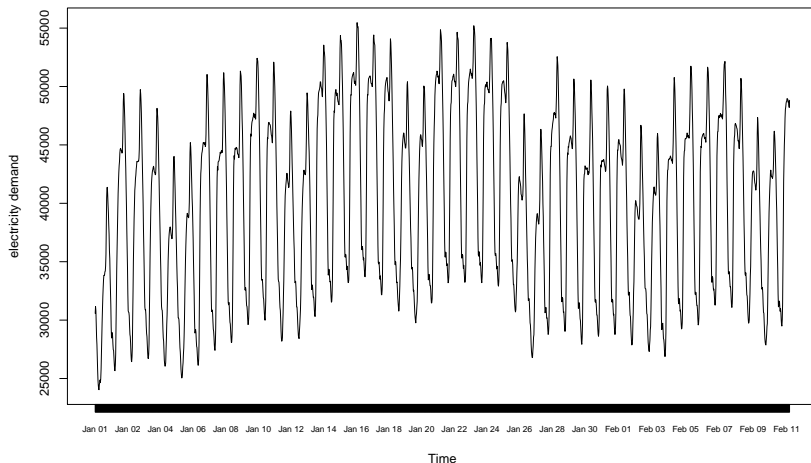
Outline

1. What is Business Analytics?

2. Overview of the Lecture

3. Organizational Stuff

Example 1: Demand Forecasting



Note: Metered half-hourly electricity demands, nationalgrid, UK, 2013.

Example 2: Recommender Systems

Your Recent History (What's this?)

Recently Viewed Items


-  [Raging Bull \(Special Edition\)](#)
DVD ~ Robert De Niro
-  [The Godfather \(Widescreen Edition\)](#) DVD ~ Marlon Brando

Top-N Items

Continue shopping Customers Who Bought Items in Your Recent History Also Bought:

-  [Chinatown](#)
(Special Collector's Edition) DVD ~ Jack Nicholson
-  [The Deer Hunter](#)
DVD ~ Robert De Niro
-  [Mean Streets \(Special Edition\)](#)
DVD ~ Julie Andreman
-  [The Graduate](#)
DVD ~ Dustin Hoffman





Selected Profile

 Marta

Number of ratings given: 12
Implicit feedbacks collected: 125
Given recommendations relevance index: 34%

Groups

Switch to a group profile:

-  **Romance**
254 users
-  **Rock**
2 134 users
-  **Alternative**
2 users
-  **Best of Nirvana**
16 users

Ratings

Predictions for you

- ★★★★★ 5.0 stars [Heat \(1995\) DVD VHS infc](#)
Action, Crime, Thriller
- ★★★★★ 5.0 stars [Professional, The \(a.k.a. I Crime, Drama, Romance, Thriller\)](#)
unique, tense, classic
- ★★★★★ 5.0 stars [Seven Samurai \(Shichinin\)](#)
Action, Drama, Japanese
- ★★★★★ No:oon [Cross of Iron \(1977\) DVD](#)
War

del.icio.us

url <http://www.bibsonomy.org/>

description BibSonomy

notes a collaborative tagging system for both bookmarks

tags folksonomy search bibtext tags

save originally posted on 2006-03-19. [delete this post.](#)

recommended tags

bibliography bibtext book bookmarking bookmarks books catalog collaborative database Folksonomies folksonomy journal library metadata Ontology papers publications research review search searchengine semantic social socialnetwork software tagging tags taxonomy technology tool tools web web2.0 world Writing

your network

for:calefeto for:frankieta for:gangelasa for:lanubile for:sarahlohnes for:teresadaniela for:zephoria

popular tags

folksonomy bookmarks bibtext tagging social tags bibliography

Example 3: Opinion Mining

★★★★★ **One of the best**, April 19, 2013

By [seen enough](#) - [See all my reviews](#)

Amazon Verified Purchase ([What's this?](#))

This review is from: [Prometheus \(DVD\)](#)

I am a **true** SCI-FI fan and this movie is in my **top 10**. The plot was **well** thought out. The chosen actors were really **well** suited for their characters. The special affects and c.g. were **top notch**. The sound quality was **very good**. The story line did get a little fragmented at times, but that happens when you have to condense a 4 hour story down into about 2 hours. Noomi Rapace was truly **OUTSTANDING** in her performance. I hope this group gets back together for another movie. I have only 2 slightly **negative** comments. The Alien creatures form and scope was so varied and inconsistent it was had to follow what came from what. And, like one reviewer said the story **left** you with more questions than answers!! This movie really should have been split into two separate movies so that more time could have been spent developing it's individual characters. **GREAT** MOVIE!!!!!!

Note: Example opinion taken from amazon.com.

Business Analytics and Machine Learning

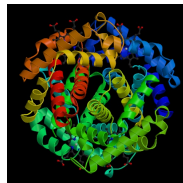
Information Systems



Robotics



Bioinformatics



**Many
Further
Applications!**

MACHINE LEARNING

OPTIMIZATION

Business Analytics and Machine Learning

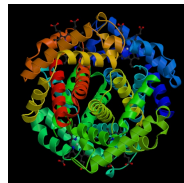
Information Systems



Robotics



Bioinformatics



**Many
Further
Applications!**

**BUSINESS
ANALYTICS**

MACHINE LEARNING

OPTIMIZATION

Business Analytics and Business Intelligence

Business Intelligence

- ▶ Business Intelligence and Business Analytics share the same goal:
to manage, process, and analyze large-scale data for supporting operations and management of businesses.
- ▶ Business Intelligence usually is focused on
 - ▶ **data preprocessing** (extract-transform-load, ETL),
 - ▶ **data management**, esp. data warehouses (DWH),
 - ▶ **reporting**, esp. Online Analytical Processing (OLAP),
- ▶ Business Analytics is focused on
 - ▶ **analyzing the data** (data mining, machine learning),
 - ▶ supporting a **specific business process**.
- ▶ Business Analytics also adopts **Big Data Analytics** that uses different data management and processing infrastructure from data warehouses.

Outline

1. What is Business Analytics?

2. Overview of the Lecture

3. Organizational Stuff

Topics of the Lecture (Current Planning)

1. Basics of Prediction
(supervised learning, regression and classification)
2. Cluster Analysis
3. Dimensionality Reduction and Visualization
4. Frequent Pattern Mining
5. Recommender Systems
6. Time Series Forecasting
7. Big Data Analytics

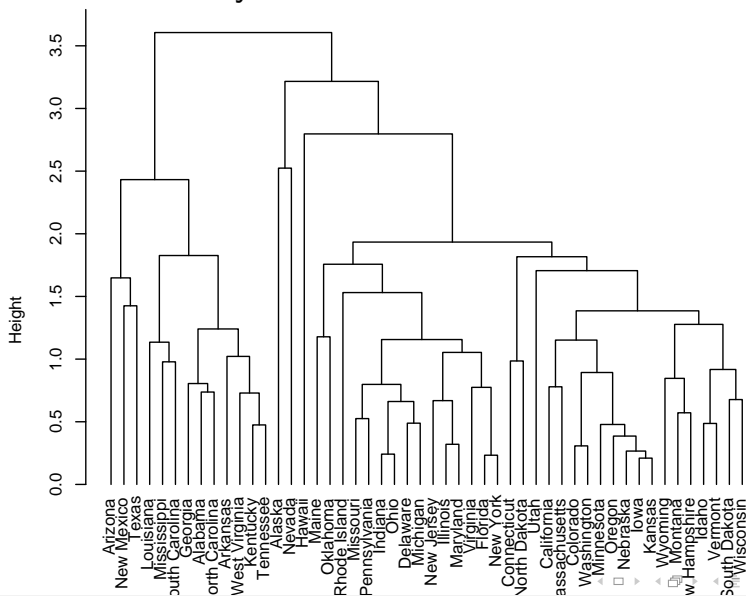
test set $\mathcal{D}^{\text{test}}$.

MAE: 21,267.44
RMSE: 28,052.88



2. Cluster Analysis

Cluster Dendrogram



3. Dimensionality Reduction and Visualization



4. Frequent Pattern Mining

Association rules in large transaction datasets:

- ▶ look for products frequently bought together (**frequent itemsets**).
- ▶ look for rules in buying behavior (**association rules**)

Examples:

- ▶ {beer, pampers, pizza} (support=0.5)
 {bread, milk} (support=0.5)
- ▶ If beer and pampers, then pizza (confidence= 0.75)
 If bread, then milk (confidence=0.75)

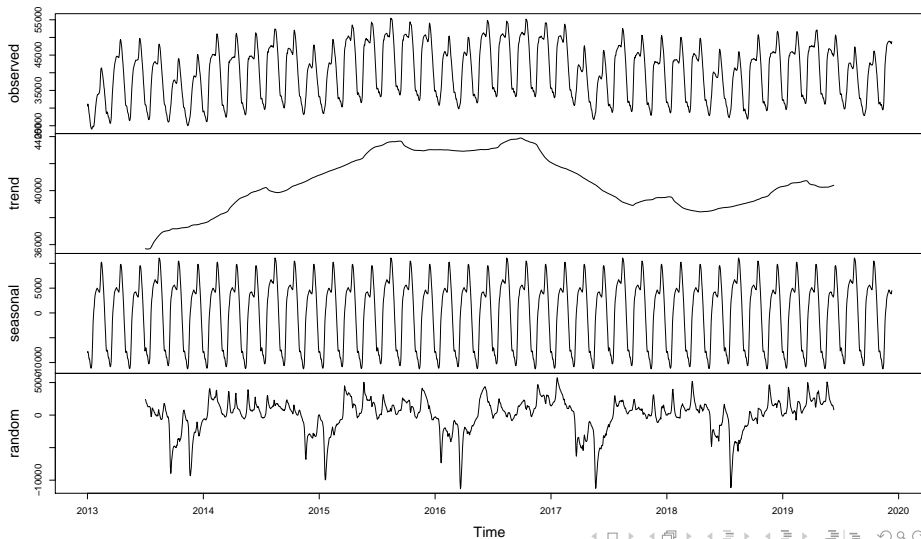
| cid | beer | bread | icecream | milk | pampers | pizza |
|-----|------|-------|----------|------|---------|-------|
| 1 | + | - | - | + | + | + |
| 2 | + | + | - | - | + | + |
| 3 | + | - | + | - | + | + |
| 4 | - | + | - | + | - | + |
| 5 | - | + | + | + | - | - |
| 6 | + | + | - | + | + | - |

5. Recommender Systems

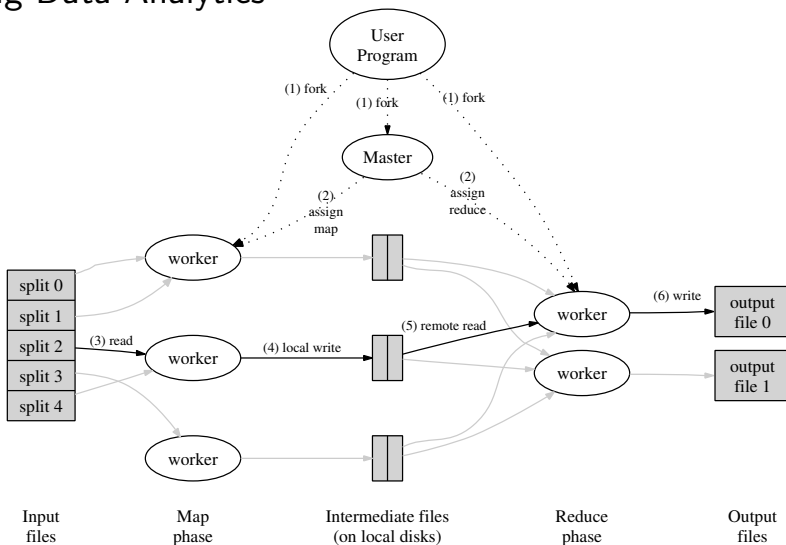
| | | | item | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ... |
|------|----------------------|------------|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | gender age ... | year | '95 | '95 | '95 | '95 | '95 | '95 | '95 | '95 | '95 | '95 | |
| | | action | - | + | - | + | - | - | - | - | - | - | |
| | | children's | + | - | - | - | - | - | - | - | + | - | |
| | | ... | | | | | | | | | | | |
| | | | item | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ... |
| user | 1 | m 24 | user 1 | 4 | | | 3 | | | | | | |
| | 2 | f 53 | 2 | 1 | | | | | | | | | |
| | 3 | m 23 | 3 | | 4 | | | | | | | | |
| | 4 | m 24 | 4 | | | | | 5 | | | | | |
| | 5 | f 33 | 5 | | | | | | | | 2 | | |
| | 6 | m 42 | 6 | | | | | | | | 4 | | |
| | : | | : | | | | | | | | | | |

6. Time Series Forecasting

Decomposition of additive time series



7. Big Data Analytics



Note: from [DG04].

Outline

1. What is Business Analytics?

2. Overview of the Lecture

3. Organizational Stuff

Exercises and Tutorials

- ▶ There will be a weekly sheet with two exercises handed out **each Tuesday** in the lecture.
1st sheet will be handed out next Tue. 30.4.
- ▶ Solutions to the exercises can be submitted until **next Monday noon**
1st sheet is due Mon. 6.5.
- ▶ Exercises will be corrected.
- ▶ Tutorials **each Tuesday 12pm–2pm**,
1st tutorial at Tue. 30.4..
- ▶ Successful participation in the tutorial gives up to 10% bonus points for the exam.

Exam and Credit Points

- ▶ There will be a written exam at end of term (2h, 4 problems).
- ▶ The course gives 8 ECTS (3+2 SWS).
- ▶ The course can be used in
 - ▶ Wirtschaftsinformatik MSc / Wirtschaftsinformatik / Gebiet BI
 - ▶ IMIT MSc. / Informatik / Gebiet KI & ML
 - ▶ as well as in both BSc programs.
 - ▶ It replaces the former separate courses “Business Intelligence” (BSc) and “Business Analytics” (MSc).

Some Books

- ▶ Wolfgang Jank (2011):
Business Analytics for Managers, Springer.
- ▶ James R. Evans (2013):
Business Analytics, Methods, Models, and Decision, Pearson.
- ▶ Mike Biere (2011):
The New Era of Enterprise Business Intelligence, Using Analytics to Achieve a Global Competitive Advantage, IBM Press.
- ▶ Daniel S. Putler, Robert E. Krider (2012):
Customer and Business Analytics, Applied Data Mining for Business Decision Making Using R, CRC Press.
- ▶ A Ohri (2012):
R for Business Analytics, Springer.

More Books

- ▶ Trevor Hastie, Robert Tibshirani, Jerome Friedman (²2009):
The Elements of Statistical Learning, Springer.
Also available online as PDF at
<http://www-stat.stanford.edu/~tibs/ElemStatLearn/>
- ▶ Kevin P. Murphy (2012):
Machine Learning, A Probabilistic Approach, MIT Press.
- ▶ Christopher M. Bishop (2007):
Pattern Recognition and Machine Learning, Springer.
- ▶ Richard O. Duda, Peter E. Hart, David G. Stork (²2001):
Pattern Classification, Springer.
- ▶ W. N. Venables, B. D. Ripley (2002):
Modern Applied Statistics with S, Springer.

Some First Machine Learning Software

- ▶ R (v3.0.0, 3.4.2013; <http://www.r-project.org>).
- ▶ Weka (v3.6.9, 22.1.2013;
<http://www.cs.waikato.ac.nz/~ml/>).
- ▶ SAS Enterprise Miner (commercially).

Public data sets:

- ▶ UCI Machine Learning Repository
(<http://www.ics.uci.edu/~mllearn/>)
- ▶ UCI Knowledge Discovery in Databases Archive
(<http://kdd.ics.uci.edu/>)

References



Jeffrey Dean and Sanjay Ghemawat.

MapReduce: simplified data processing on large clusters.

In *OSDI'04 Proceedings of the 6th conference on Symposium on Operating Systems Design & Implementation*, volume 6, 2004.