MTAT.03.319

Business Data Analytics

Lecture 1: Introduction

Rajesh Sharma

Marlon Dumas will give the last lecture on BPM

Slides: Thanks to Anna and Marlon

FirstName.LastName@ut.ee
Your background

Your expectations
Warm-up question

• We are a charity. We have a database containing 100K donors who have not donated in the past 12 months. We know their basic demographics, address and how much they have donated in past (and when). We do not have emails.

• Sending a mail asking for donation costs 60 cents/piece. When we mail out, the average donation comes at about 80 cents.

• Should we send a (postal) mail to all 100K donors?
What is Business (Data) Analytics?
Business Analysis vs Business Data Analytics
BA Vs. BDA

• **Business Analysis**
  
  • Analyzing the problem and mapping it to the solution space (without the data)
  
  • Qualitative in nature
  
  • When data is not available or cannot be get (time constraint)
  
  • Eg: Put adv on street for donation.

• **Business DATA Analytics**
  
  • Mapping of the problem to the solution space based on the data analysis
  
  • Quantitative in nature
  
  • Take actions based on the data.
  
  • Eg: Send mails to people who have donated in the past based on analysis.
The application of repeatable methods to use data in order to inform or make decisions to achieve or maximize a business objective.
Business Intelligence vs Business (Data) Analytics
Business Intelligence vs. Analytics

• Do 18-30 y.o. male customers living in cities of 50K+ inhabitants buy more of product X than others?

• In which zip codes have the sales of product X increased the most in the months of June-July?

• If a customer buys product X, which other products might he/she buy?

• Are customers whose number of calls in the past 3 months is lower than the 12-months average more likely to churn in the next 6 months?

• Which customers are most likely to buy product X?

• How much will we sell of product X in each zip code next month?

• For each customer, which product are they most likely to buy given their current shopping cart?

• ???
Typical classes of analytics questions

• Clustering: Which objects belong together/are similar?

• Classification: Who likes X? Who will do X? …

• Regression: How much will X spend/buy/earn/donate …?  

• Censored regression: How much will X spend/buy/earn/donate… knowing that they might not spend/buy/earn/donate at all…

• Forecasting: What will be the price/cost of X in future? What will be the average expenditure/income, etc. of a population in future?

• …
Business Intelligence vs. Analytics


Data collection & preparation (auto) → Decision Model Construction (auto) → Action
Exercise

• I am a marketing manager in a company that sells organic products online. I have demographic and purchasing data about every single customer - every single purchase they’ve made, every single visit to our web site, durations, pages visited, clicks they’ve made to our past email campaigns, etc.

• I can send 100K Christmas electronic flyers via email next month. My flyer design team offers me to produce up to six different flyers. I can choose the products included in each flyer and the layout of the products.

• I want to maximize the revenue resulting from this campaign. I don’t care about stock right now – assume we have infinite amounts of every product.

• I am fortunate enough to have a data analyst “guru” for one week. What should I ask him for?

• Approach this question using data Intelligence vs. data analytics?
Types of Business Analytics

Descriptive
- What has happened?
  - E.g. what top five customer segments we have?
  - Which pairs of products are bought together?

Predictive
- What will happen?
  - E.g. Who will buy?
  - Who will churn?

Prescriptive
- What to do to achieve my goals?
  - When should I make my next customer call, to whom and what should I tell them?
Other terms

• Business Intelligence
  • Dashboard and report extraction from enterprise databases for monitoring, data exploration and hypothesis validation

• (Statistical) Data Analysis
  • Very useful for exploration & hypothesis validation; works with “small data”; driven by models

• Data Mining: Extracting patterns from data automatically
  • Business analytics applies data mining to business problems

• Machine learning
  • The algorithms behind data mining

• Data Science
  • Catch-all term for data analysis, analytics, mining, machine learning, deep learning, Big Data processing, network science and other techniques
When & where to use business analytics?
## Business problems

<table>
<thead>
<tr>
<th>CRM</th>
<th>BPM</th>
<th>BRM</th>
</tr>
</thead>
</table>
## Application Areas of Business Analytics

<table>
<thead>
<tr>
<th>CRM</th>
<th>BPM and ERP</th>
<th>BRM</th>
</tr>
</thead>
</table>

- Marketing
- Sales
- **Customer Lifecycle Management (CLM)**
- Process optimization
- Inventory management
- HR/Talent management
- Facilities management
- Risk management
- Underwriting
BA in CRM

Customer segmentation:
- Behavioral
- Product/category based
- Brand based

Product Recommendations
- Up sell
- Cross sell
- Next sell

Customer Lifecycle Management
- Acquisition
- Retention
- Win-back
Figure 2-1: The customer lifecycle progresses through different stages.
BA in BPM

• Discover how a process is actually executed
  • Including deviations with respect to how it should be done
• Identify root causes of poor process performance
  • E.g. customer complains, defects, rework, delays, etc.
• Predict that a given execution of a process will end up in an undesirable outcome
  • E.g. predict delays at runtime, before they happen
BA in BRM

Strategic and compliance
- industry changes?
- economic stability
- merger and acquisition activity

Financial and operational
- recruitment
- supply chain
- fraud detection
- investments, portfolio
- credit scoring
How business analytics works?
CRISP-DM
Cross-Industry Standard Process for Data Mining

Define the project

- Examine the data; identify issues with data
- Fix data issues; create derived variables (features)
- Build models
- Assess the goodness of the model and their expected impact

Put models and insights into use

- Deployment
- Evaluation
- Data Preparation
- Modeling
- Data Understanding
- Business Understanding

Data
Who is involved?

- Business sponsor
- Domain expert(s)
- Analytics expert
- Data steward & database expert
Business understanding

• Define the business objective
• Formulate the question(s)
• Identify target variable
• Identify unit of analysis
• Identify attributes (also called *features*)
Back to warm-up question

• We are a charity. We have a database containing 100K donors who have not donated in the past 12 months. We know their basic demographics, address and how much they have donated in past (and when). We do not have emails.

• Sending a mail asking for donation costs 60 cents/piece. When we mail out, the average donation comes at about 80 cents.

• Should we send a (postal) mail to all 100K donors?
CRISP-DM & Course Structure

Lecture 1
- Business Understanding
- Data Understanding
- Deployment
- Evaluation

Data

Lecture 2
- Data Preparation
- Modeling

Lecture 3

Lectures 4-8

Out of scope
Lecture Topics

1. Introduction
2. Descriptive analysis and visualization
3. Customer segmentation
4. Customer Lifecycle management - regression problems
5. Customer Lifecycle management - classification problems
6. Models Classification
7. Cross-sell/Up-sell recommendations: Market Basket Analysis
8. Cross-sell/Up-sell recommendations: Collaborative Filtering
9. A/B testing in marketing
10. Financial Forecasting
11. Guest Lecture by Veronika Plotnikova, Swedbank: Topic and exact date to be decided
12. Link Analysis 1: Introduction and Basics about Network Science
13. Link Analysis 2: Use cases -- Telecomm, Companies, Social Capital ..
14. Social Media Analysis: Mining and Understanding Text.
15. Business Process Mining, Lecture by Marlon Dumas
Course Structure & Assessment

Class participation: At least 2 exercises every week
5 points each
Students will be picked randomly from the class. They have to demonstrate the solution for the exercises given last week. Every student will be called 2 times in the entire course.

4 homeworks. 10 points each
deadline = At 24.00 before next lecture, -1 p for each day late
Individual or in pair (2)

1 exam
50 points, 20 is minimum to pass the course
Theory + Practical
Logistics

• Meeting hours

  • Room 119 Juhan Liivi 2, Tartu.
  
  • Time: 10.30 to 13 (every Tuesdays)

• Laptop:

  • You should have a laptop as you will be called randomly and need to show the solution of the previous lecture’s exercises
  
  • In case you need laptop, contact librarian at the 5th floor in Liivi 2.
Something about you?

- How many of you know R?
- Background: 1) Business 2) Computers/IT and 3) Non Technical
- Compulsory and Non Compulsory
- Full time student Vs. Open University
- Business Department Vs. Non Business Background
Course information

all the relevant information appears on courses webpage:


Warm-up homework (no points, no submission)

https://courses.cs.ut.ee/2018/bda/spring/Main/Homework0
NOTE

• It is going to be an interactive class
  • Problem oriented course
  • Theory
  • Practical

• Ask as many questions you want. Do not be shy.

• If I cannot answer, I will try to google it. I will not be shy!

• Discussions through Piazza: No anonymous posts: You are welcome to criticize openly **but** constructively.
R is a **language** and environment for statistical computing and graphics.

R is an open source.
RStudio

https://www.rstudio.com/

is a graphical development environment
Homework

See course Web page
No points given, nothing to submit, it’s for you to prepare!