Lecture 2
–
Business Understanding

Blaumeise / Blue Tit (*Cyanistes caeruleus*)

- The blue tit is a small bird in the tit family. It is easily recognisable by its blue and yellow plumage and small size.
- Blue tits, usually resident and non-migratory birds, are widespread and a common resident breeder throughout temperate and subarctic Europe and the western Palearctic in deciduous or mixed woodlands with a high proportion of oak.
- They usually nest in tree holes, although they easily adapt to nest boxes where necessary.
- The blue tit prefers insects and spiders for its diet. Outside the breeding season, they also eat seeds and other vegetable-based foods.
- The birds are famed for their acrobatic skills, as they can hold on to the outermost branches of trees and shrubs and hang upside down when looking for food.

Sources:
* Photo by federicomaderno on pixabay
* https://en.wikipedia.org/wiki/Eurasian_blue_tit

Business Understanding

1. CRISP-DM: Business Understanding
2. Fundamental Data Science (ML) Tasks
3. Business Problem Decomposition

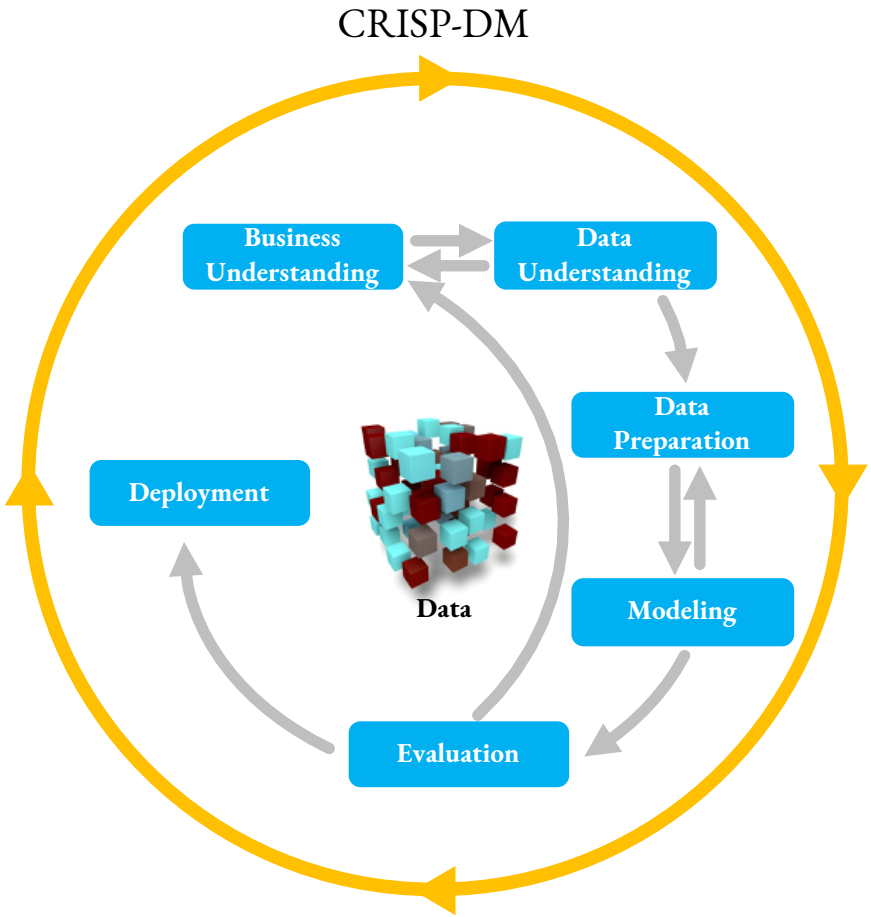
Business Understanding Phase of CRISP-DM

Business Understanding

- Determine business objectives
- Assess situation
- Determine data mining goals
- Produce project plan

not covered in this class

➔ Extremely critical phase



CRISP-DM: Business Understanding

- Determine business objectives
  - Describe the customer's primary objective, from a business perspective. In addition to the primary business objective, there are typically other related business questions that the customer would like to address.
- Assess situation
  - This task involves more detailed fact-finding about all of the resources, constraints, assumptions and other factors that should be considered in determining the data analysis goal and project plan. In the previous task, your objective is to quickly get to the crux of the situation. Here, you want to flesh out the details.
- Determine data mining goals
  - A business goal states objectives in business terminology. A data mining goal states project objectives in technical terms.
- Produce project plan (→ not covered in this class)
  - Describe the intended plan for achieving the data mining goals and thereby achieving the business goals.

From Business Objective and Situation to the Business Problem

- Business Problem needs to be solved to achive the Business Objective in the given Situation
  - Stated in the language of the business domain
  - Usually not solvable directly
  - Multiple Business Problems possible and quite common

- Examples
  - Car repair shop: “Customers bring in their cars and ask us to refill their oil. Frequently, they do not know which oil they put in last time, but we need this information in order to refill with the correct oil.”
    - Business Objective: Refill with the correct oil.
    - Business Problem: Figure out which oil was used last and currently is in the car.

  - Mobile phone operator: “Customer acquisition costs are very high. We need to reduce the number of customers leaving at the end of their contract.”
    - Business Objective: Increase number of customers.
    - Business Problem: Reduce number of customers leaving at the end of their contract.

Demo

Demo

Business Objectives and Business Problems

Exercise

Exercise 1

Business Objectives and Business Problems

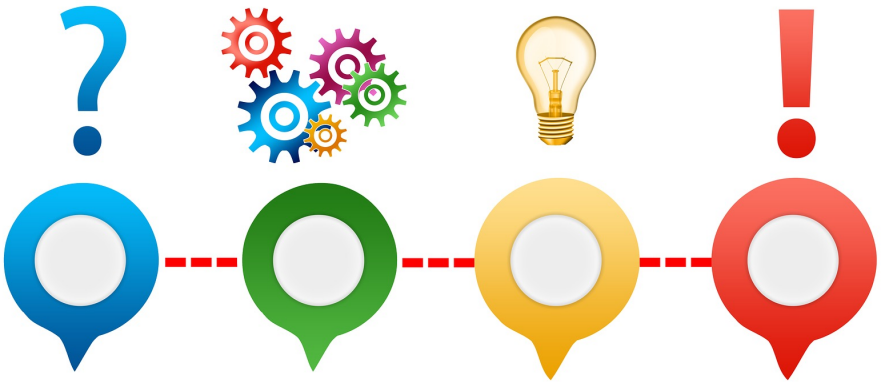


Photo by Gerd Altmann on Pixabay

Business Understanding

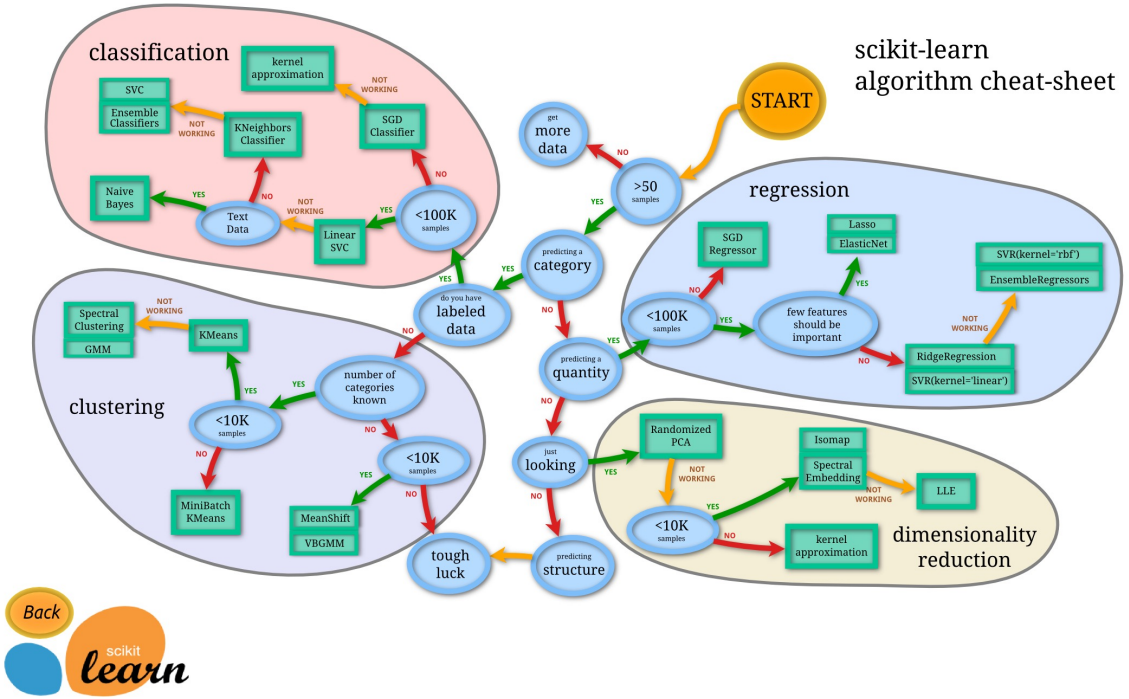
1. CRISP-DM: Business Understanding
2. Fundamental Data Science (ML) Tasks
3. Business Problem Decomposition

Decomposition of Business Problem into ML Problem

„A critical skill in data science is the ability to decompose a data-analytics problem into pieces such that each piece matches a known task for which tools are available.“ (Provost/Fawcett)

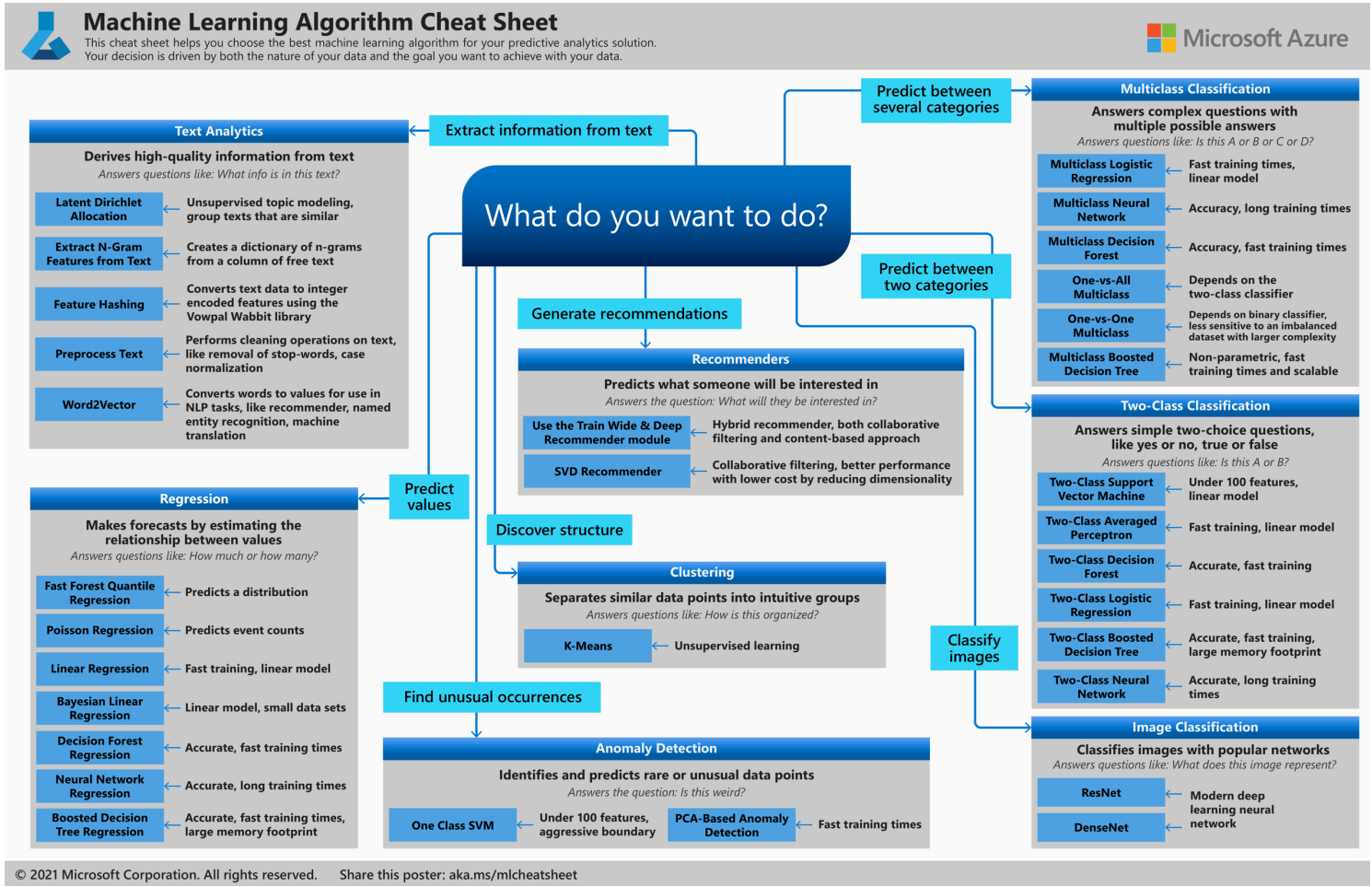
Fundamental (Data Science/ML) Tasks

- Numerous fundamental Tasks
- Usually, multiple algorithms for each task
- Examples
  - Classification and class probability estimation
  - Regression
  - Clustering
  - Anomaly Detection and Profiling
  - Link Prediction
  - …



Source: https://scikit-learn.org/stable/tutorial/machine_learning_map/index.html

Machine Learning Tasks and Algorithms - Microsoft



Source: https://docs.microsoft.com/en-us/azure/machine-learning/algorithm-cheat-sheet

Business Understanding

1. CRISP-DM: Business Understanding
2. Fundamental Data Science (ML) Tasks
3. Business Problem Decomposition

Business Problem Decomposition: From Business Problem to Data Science Tasks

- Business Domain
  - State the Business Objective and the Business Problem in the language of the business domain
  - Usually not solvable directly
- Data Science Problem
  - Stated the Data Science Idea
  - Decompose Business Problem into fundamental Data Science tasks
  - List the data needed as input for the chosen tasks
  - Frequently, multiple approaches/decompositions possible

Business Problem Decomposition: Example

- Business Domain
  - Car repair shop: “Customers bring in their cars and ask us to refill their oil. Frequently, they do not know which oil they put in last time, but we need this information in order to refill with the correct oil.”
    - Business Objective: Refill with the correct oil.
    - Business Problem: Figure out which oil was used last and currently is in the car.

- Data Science Problem
  - Data Science Idea
    - Use a spectrometer to analyse the chemical composition of the most frequently used oils
    - When a car comes in, extract a sample of the oil, analyse the chemical composition with the spectrometer and find the closest match
    - If a close match has been found, assume that this is the oil in the car
  - Fundamental Data Science/ML Task
    - Classification and Class Probability Estimation to evaluate the closeness of the match
  - Data needed
    - Acquire enough spectrographic analysis
    - Select best features for classification
    - Decide on the threshold for class probability estimation

Demo –Business Problem Decomposition

Business Domain: Cement Production

The ENCI cement production factory (Maastricht)

Source: https://www.enci.nl/nl/enci-maastricht

Demo

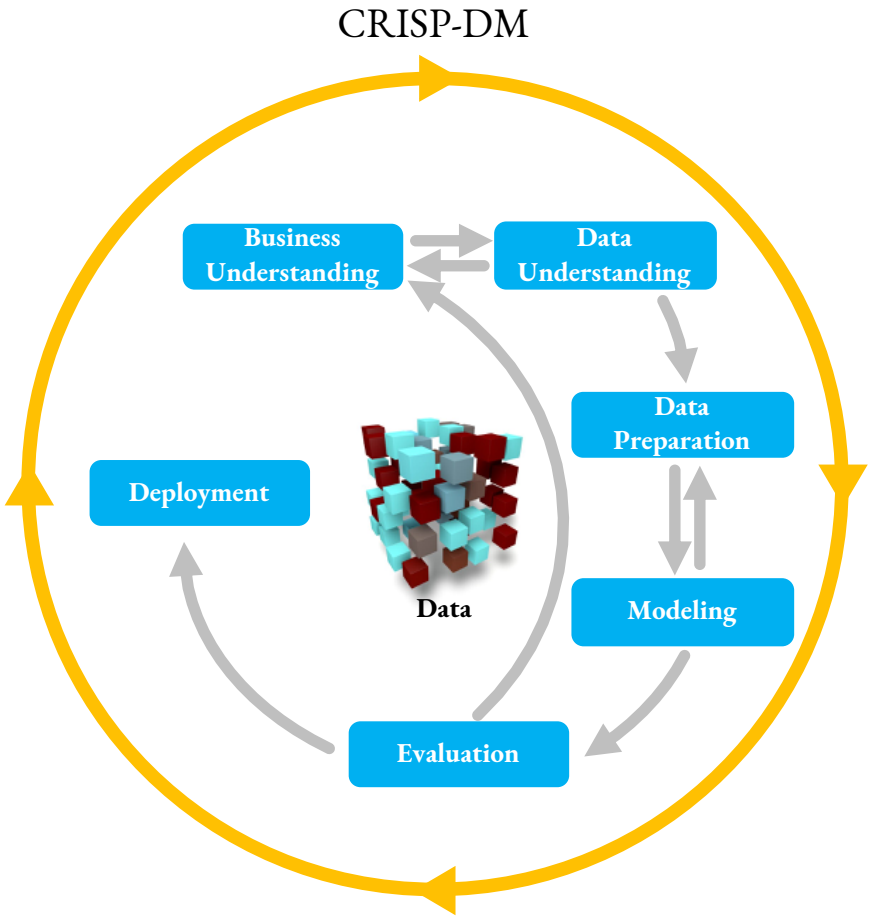
Demo

Business Problem Decomposition

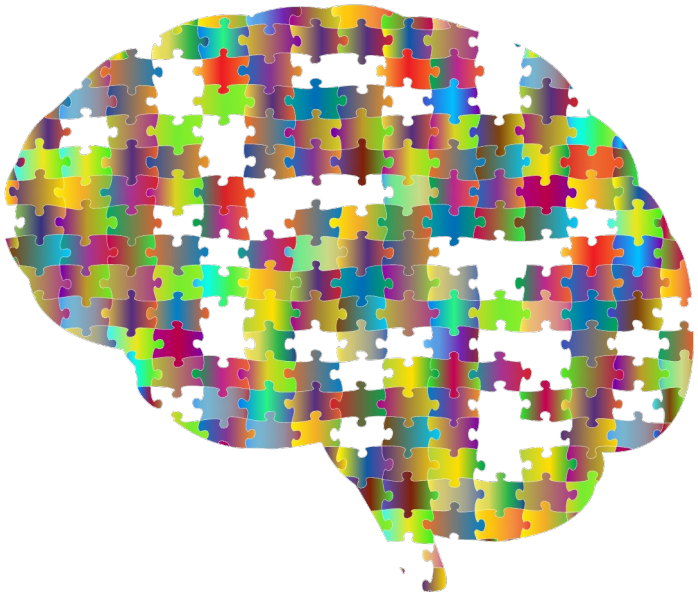
Iterative Approach

Common Iterations

- Iterate from from Data Understanding
- Iterate back from Evaluation
- Data Science Problem and Tasks may change
- Business Problem may change
- Business Objective may change



Key Takeaways



- Fundamental Data Science (ML) Tasks
- Four Parts of the Business Understanding Phase of CRISP-DM
- Business Domain
  - State Business Objective
  - State Business Problem
- Data Science Problem
  - State Data Science Idea
  - Decompose Business Problem into fundamental Data Science Tasks
  - State Data needed

Image by Gordon Johnson on pixabay

Exercise

Exercise 2

Business Problem Decomposition

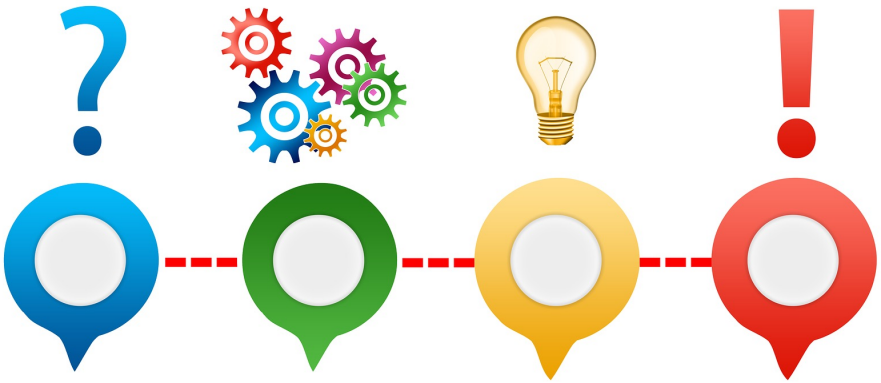


Photo by Gerd Altmann on Pixabay