Data Science for Business

Witek ten hove

2020-12-03

Contents

Preface

\begin{figure}

{

}

\caption{CRISP-DM Model taken from: https://commons.wikimedia.org/wiki/File:CRISP-DM_Process_Diagram.png}
\end{figure} ## Prerequisites {-}

Before starting this module make sure you have:

- access to the book *Provost, F., & Fawcett, T. (2013). Data Science for Business: What you need to know about data mining and data-analytic thinking. O'Reilly Media, Inc.*
- installed R and RStudio
- a Github account

Purpose of this course

The general learning outcome of this course is:

> The student is able to perform a well-defined task independently in a relatively clearly arranged situation, or is able to perform in a complex and unpredictable situation under supervision.

The course will provide you with a non-technical overview of data science, and types of techniques. The focus will lie on critical thinking and the full DS process (based on CRISP).

Structure of the course

Table 1: Course overview

Week nr.	Module name	Readings
1	Onboarding and introduction to course	
2-3	Data-Analytic Thinking	Provost / Fawcett Ch.1
4-5	Business Problems and Data Science Solutions	Provost / Fawcett Ch.2
6-7	Data Science and Business Strategy	Provost / Fawcett Ch.13

About the author

Witek ten Hove is a senior instructor and researcher at HAN University of Applied Sciences. His main areas of expertise are Data en Web Technologies.

Through his extensive business experience in Finance and International Trade and thorough knowledge of modern data technologies, he is able to make connections between technology and business. As an open source evangelist he firmly believe in the power of knowledge sharing. His mission is to inspire business professionals and help them exploit the full potential of smart technologies.

Chapter 1

Data-Analytic Thinking

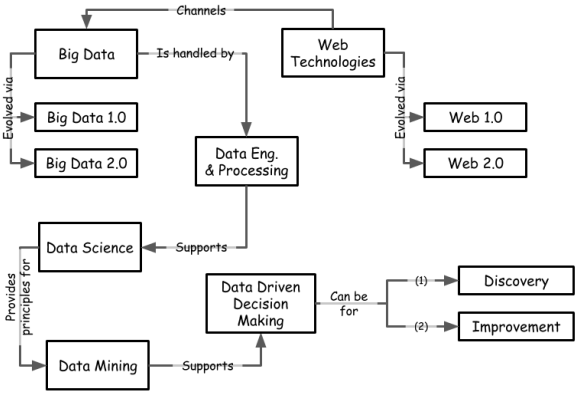


Figure 1.1: Lesson Map (own work)

1.1 What is Data Science and what is not?

What is the ultimate purpose of a Data Science effort?

1.2 What is Data Mining?

\begin{figure}

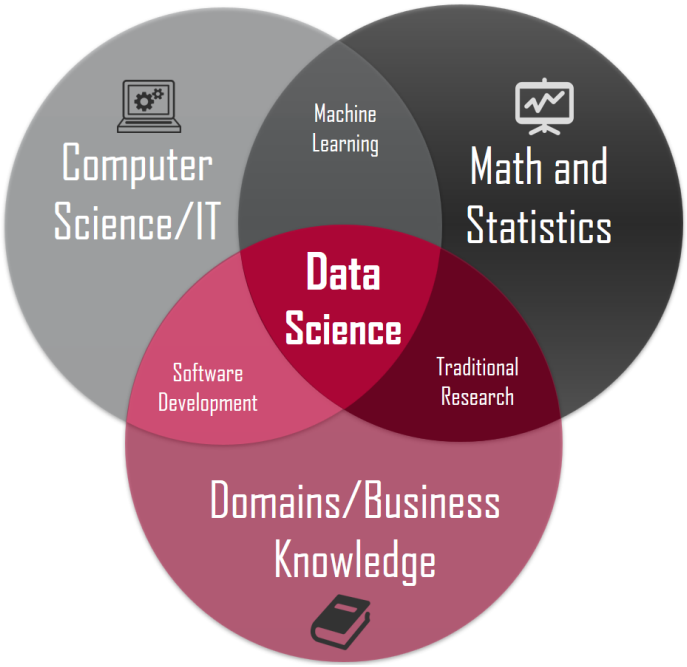


Figure 1.2: Data Science diagram taken from: https://towardsdatascience.com/introduction-to-statistics-e9d72d818745

{

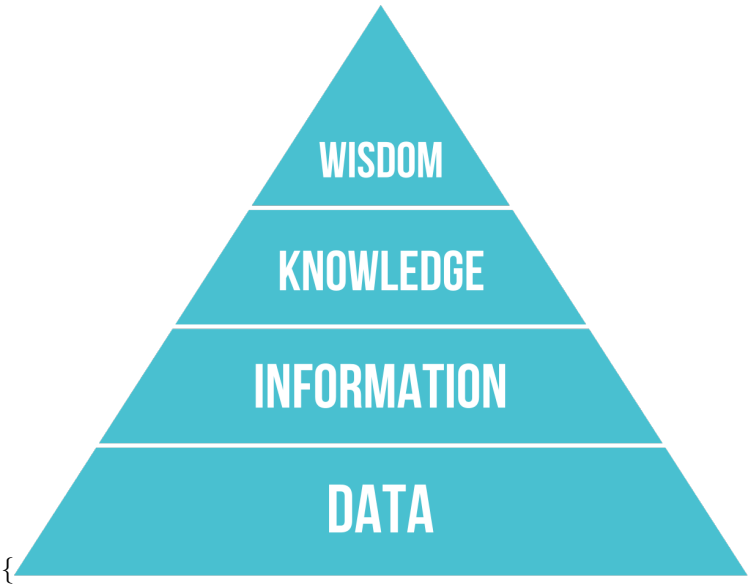
}

\caption{Cross-industry standard process for data mining taken from: https://commons.wikimedia.org/wiki/File:CRISP-DM_Process_Diagram.png}
\end{figure}

1.3 What is Data Driven Decision Making

What are other types of decision making? (relation with What is proof or evidence?)

\begin{figure}



{

}

\caption{DIKW pyramid taken from: https://en.wikipedia.org/wiki/DIKW_pyramid#/media/File:DIKW_Pyramid.svg}
\end{figure}

1.4 What are the challenges and opportunities that Big Data creates?

1.5 What are Big Data technologies?

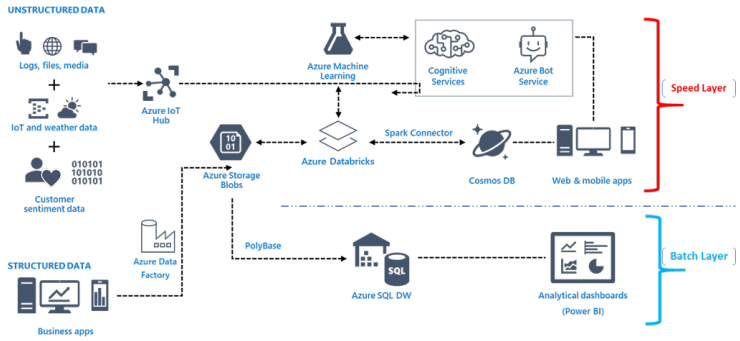


Figure 1.3: Big Data pipelines taken from: https://azure.microsoft.com/en-gb/blog/the-emerging-big-data-architectural-pattern/

1.6 Case Study: Ubernomics

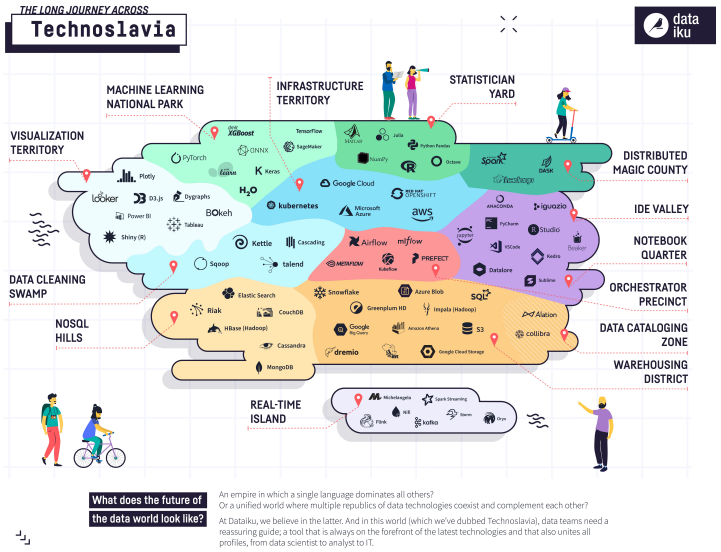


Figure 1.4: Big Data pipelines taken from: https://blog.dataiku.com/technoslavia-the-fragmented-world-of-data-infrastructure-in-2020

Chapter 2

Business Problems and Data Science Solutions

Formulating a business problem using ECLIPSE (Wildridge and Bell, 2002)

Table 2.1: ECLIPSE questions

Mnemonic	Description
E	Expectation: Why does the user want the information?
C	C - Client Group: For whom is the service intended?
L	L – Location: Where is the service physically sited?
I	Impact: What would represent success? How is this measured?
P	P – Professionals: Who provides or improves the service?
SE	SE – Service: What type of service is under consideration?

Chapter 3

Data Science and Business Strategy

3.1 Case Study: AI Business

Chapter 4

Applications

Some *significant* applications are demonstrated in this chapter.

4.1 Example one

4.2 Example two

Chapter 5

Final Words

We have finished a nice book.

Bibliography

Wildridge, V. and Bell, L. (2002). How clip became eclipse: a mnemonic to assist in searching for health policy/management information. *Health Information & Libraries Journal*, 19(2):113–115.