A Journey to Data Scientist

(Cécile Bothorel)

This course offers hands-on activities for mining datasets, including the use of descriptive statistics tools and the main data mining algorithms. It draws on the CRISP-DM data analysis methodology to identify and articulate the different phases of data analysis, starting with the definition of a business need. We'll also look at how to validate and present the results to decision-makers. We'll provide theoretical contributions, especially in law and economics, to help you make the most of data while respecting legal issues. Plus, we'll do some practical activities that cover the fundamental principles of effective data visualization and storytelling techniques to help you enhance data-driven decision-making.

This course is all about putting theory into practice. It's based on math and computer science, but it also covers legal, economic, and storytelling with data. You'll get to put all three of these areas into practice through a "fil rouge" project.

Key concepts covered:

- CRISP-DM methodology for data analysis
- Business problem definition and validation criteria
- Data economics and valorization strategies
- Ethical and legal aspects of data usage
- How and when use descriptive statistics and principal data mining algorithms
- Principles of effective data visualization
- Data storytelling techniques and best practices

By the end of this course, students will be able to:

- [BC-01] Conduct a data science project in line with business objectives and considering ethical and legal implications
- [BC-03] Define and analyze relevant business problems using data science techniques
- [BC-04] Design and create impactful data visualizations and stories to enhance decision-making processes
- [BC-07] Apply the CRISP-DM methodology to real-world data analysis projects

Commenté [BC1]: BC-01 ou 02 ?

Prerequisites:

- Being familiar and efficient with Python programming
- Being familiar with key libraries for Data Science like Pandas, NumPy, scikit-learn