

2024-2025

YEAR 2

**BIG DATA & MANAGEMENT** 

PROGRAMME SUPERVISOR(S) :

Luisa ROCHA DA SILVA

Marc GIBIAT



BBA – Big data & Management

# YEAR 2 – Autumn Semester

## **CORE COURSES**

Course code	Title	ECTS Credits
BUSETHICS	Business ethics	3
CONSUM	Consumer study	3
EPSIII	Sports	2
INFOSYS1	Information systems management 1	3
INFOSYS2	Information systems management 2	3
INFOV	Databases	3
INFOVI	Operating systems	3
MATHSV	Mathemetics V : probabilities and statistics	3
NEGO	Negotiation	3

### LANGUAGE COURSES

Course code	Title	ECTS Credits
FLEIII	French	4
LV2II	Chinese	4



BBA – Big data & Management

# YEAR 2 – Spring Semester

## **CORE COURSES**

Course code	Title	ECTS Credits
AI	Artificial intelligence fundamentals	3
AI4BUS	AI and data mining for digital business	3
CLEAN	Data cleansing and data quality	3
CLOUD	Cloud computing	3
DATAANAL	Data analysis	3
DATAMIN	Data mining	3
MLDEEPL	Machine and deep learning	3
NLP	Natural Language Processing	3
SCIVIZ	Scientific visualization	3
SSI	Systems security information	3





2024/2025 - Year 2 - Autumn Semester

# Business ethics [BDM\_BUSETHICS]

LEAD PROFESSOR(S): Luisa ROCHA DA SILVA

#### Requirements

#### Objectives

Identify and explain ethical issues that might arise in organizational settings Imagine why decision makers might choose certain actions over others when faced with an ethical issue Reflect on your own values and how they influence your ethical choices Compare different ethical theories and understand how they would impact decision making

#### **Course contents**

Students will contribute to a group project that they will start the semester with worth 40% of the final grade. Each group will choose an ethical

theory that they will then teach to the class. The lesson plan should be dynamic, creative, and encourage student participation.

At the end of the semester, there will also be an exam worth 60% of the final grade. For the exam, students will be asked to write short essays in

which they will analyze complex situations (LO2) and identify ethical issues and how to act on them from an ethical perspective.

#### Course material

Fisher, C., Lovell, A., Valero-Silva, A. (2013). Business ethics and values. Pearson. Stanwick, P. A., & Stanwick, S. D. (2021). Absolute essentials of business ethics. Routledge.

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Autumn Semester

# Consumer study [BDM\_CONSUM]

LEAD PROFESSOR(S): Luisa ROCHA DA SILVA

#### Requirements

#### Objectives

At the end of this course, the student will:

- Understand the difference between qualitative and quantitative methods of inquiry
- Learn how to develop surveys
- Develop analytical skills
- Understand how to carry out a research project
- Appreciate academic literature and scientific research
- Application of qualitative and quantitative analysis

#### **Course contents**

- Consumer Behavior essentials
- Quantitative & qualitative research
- Sampling methods and development of surveys
- Construction of surveys with Qualtrics and research ethics
- Analysis of Customer experience through text analytics
- Testing hypotheses in quantitative research through SPSS
- Analysis of visual data
- The hybrid human-machine intelligence approach to study complex phenomena

#### Course material

Hair, J. F., Page, M., & Brunsveld, N. (2019). Essentials of business research methods. Routledge. Softwares: SPSS – Leximancer - NVIVO

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Autumn Semester

# Sports [BDM\_EPSIII]

LEAD PROFESSOR(S): Gildas GUIHENEUF-LALERE

#### Requirements

#### Objectives

Sport:

- as a compensatory element, a factor of physical and moral health (general physical preparation)
- as a unifying element of cohesion through collective activities that develop the notions of cooperation, opposition, mutual aid, confidence, etc.
- as a factor of autonomy and equilibrium
- as a support for role-playing on themes linked to leadership and group management.

#### **Course contents**

- Take an active role in one's training by choosing one's sport
- Understand and assimilate the physiology of effort
- Design, evaluate and regulate learning situations
- Organize the work and production of others
- Take charge of a group in a learning situation:
  - define the context
  - define the objectives to be achieved
  - define the methods of execution
  - answer team members' questions
  - be capable of self-assessment

#### **Course material**

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	2	0 hrs	0 hrs	16 hrs	0 hrs	0 hrs





2024/2025 - Year 2 - Autumn Semester

# Information systems management 1 [BDM\_INFOSYS1]

LEAD PROFESSOR(S): Luisa ROCHA DA SILVA

#### Requirements

#### Objectives

The objective of this course is to provide a general culture in information systems, transformation and technology. It should allow students to identify the major issues that technological development takes on and give them the knowledge bases necessary for understanding more specialized courses.

#### Course contents

Theme 1: development and evolution of SI

Over the past 30 years, companies have used information systems to accelerate the automation of their business processes, improve them and transform them. These steps required significant changes that are still not stabilized, but which rely on a hard core that we will study in this first part: the information system and data management.

Theme 2: Evolution of technologies

Today, beyond the issues related to the implementation and use of information systems, the acceleration of the evolution of technologies also impacts organizations. The emergence of big data has required the implementation of new methods that can be based in particular on artificial intelligence. This presents challenges but also unprecedented opportunities for organizations that still need to be fully understood. We will use the concept of machine learning to identify how artificial intelligence can improve certain processes.

Theme 3: Social and environmental implications of technologies

The use of technologies has a strong impact on society and the environment, negatively but also positively. This third theme will be devoted to the identification of these impacts. We will deal more particularly with the notion of digital sobriety, in order to encourage companies to have a more responsible approach to their tools and their data.

#### **Course material**

#### Assessment

Individual assessment: EVI 1 (coefficient 1)							
LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM	
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs	





2024/2025 - Year 2 - Autumn Semester

# Information systems management 2 [BDM\_INFOSYS2]

LEAD PROFESSOR(S): Luisa ROCHA DA SILVA

#### Requirements

#### Objectives

The objective of this course is to show students how to use raw data to do strategic analysis, using a tool present in all companies, Excel.

Three tools, present in Excel, are Business Intelligence tools:

1. Power Query is a tool that facilitates and automates data manipulation.

2. Power Pivot allows you to create a structured data model in Excel and to increase the analytical capacities of pivot tables tenfold thanks to the DAX language.

3. Choropleth Maps and 3D Maps allow you to work with geospatial data.

It is mandatory that students work on a Windows environment. This implies that students with a MAC install a Windows desktop.

#### **Course contents**

Business intelligence (BI) or decision support system (DSS) is information technology used by decision-makers and company managers. It refers to the means, tools and methods used to collect, consolidate, model and restore data, whether tangible or intangible, from a company in order to provide decision support and allow a decision-maker to have an overall view of the activity being processed.

This type of application is based on a common architecture.

1. Operational data is periodically extracted from heterogeneous sources: flat files, Excel files, databases (DB2, Oracle, SQL Server, etc.), web services, massive data and stored data in a data warehouse.

2. The data is restructured, enriched, aggregated, reformatted, and nomenclatured to be presented to the user in a semantic form (meaningful business views) that allows decision-makers to interact with the data without having to know its physical storage structure, other schemas that allow facts and measures to be distributed according to hierarchical dimensions, preprepared reports that can be parameterized, and more synthetic and interactive dashboards.

3. This data is delivered to the various functional areas (strategic management, finance, production, accounting, human resources, etc.) through a security system or specialized datamart for consultation, analysis, pre-defined alerts, data mining, etc.

Business Intelligence is therefore the set of tools and methods aimed at transmitting relevant information to company managers. Its goal is to help them understand their environment and to support them in their strategic decision-making. The user's needs are therefore at the heart of effective Business Intelligence.

Originally, Business Intelligence was the prerogative of a few experts and the departments concerned could be counted on the fingers of one hand: accounting, finance, logistics, possibly commerce, etc. However, for several years now, web technologies and the increase in computing power of IT tools have facilitated the distribution of BI tools to users, in discrete forms. Most businesses now use Business Intelligence to consolidate the information they have at their disposal.

#### **Course material**

Business Analytics : a management approach. R. Vidgen, S. Kirshner & F. Tan. Bloomsbury Academic, 2022 Business Intelligence avec Excel : des données brutes à l'analyse stratégique. Boris Nord. ENI Editions, 2020



Business Intelligence simple et efficace avec Excel et PowerPivot. Jean-Philippe Gouigoux, 2014 Analyser efficacement vos données à l'aide des tableaux croisés dynamiques. Pierre Rigollet. ENI Editions, 2020 Datavisualisation : utiliser le storytelling pour faire parler vos données. Cole Nussbaumer Knaflic. Pearson, 2019

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Autumn Semester

# Databases [BDM\_INFOV]

LEAD PROFESSOR(S): Lucas LESTANDI

#### Requirements

#### Objectives

Understand the basics of databases (DB) construction with relational modeling, be able to create and interact with SQL driven DB and introduction to modern distributed DB architectures (noSQL).

#### **Course contents**

- Represent real world problems with DB compatible models such as relational model
- Extract information from existing DB using SQL and python APIs
- Data-mining and automatic generation of reports in the context of BI
- Creation and management of a database
- True big data with distributed DBs (noSQL)

#### Course material

Bases de données 2è édition, J-I Hainaut, Dunod, 2012

#### Assessment

Collective assessment: EVC 1 (coefficient 0.3)

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	12 hrs	0 hrs	12 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Autumn Semester

# Operating systems [BDM\_INFOVI]

LEAD PROFESSOR(S): Vincent TOURRE

#### Requirements

#### Objectives

The objectives of this course are to help students understand what Information Systems are, the stakes of having a welldesigned Information System and the challenges associated with it. Students will learn how to use modelling tools such as SysML and BPMN, which are essential to specification and design of Information Systems. Then, an overview of Enterprise Architecture will be given, as a method to represent the current state of an Information System, plan for future transformations and inform business decisions.

#### **Course contents**

- Introduction to Information Systems
- Using SysML to model the different aspects of a business and support communication between the various professional bodies inside a company
- Modelling business processes with BPMN
- Overview of the concepts of Enterprise Architecture

#### Course material

- C. Longépé, Le Projet d'Urbanisation du SI, Dunod, 2009
- S. Friedenthal, A. Moore and R. Steiner, A Practical Guide to SysML, The MK/OMG Press, 2015
- M. Dumas, Ma. La Rosa, J Mendling, H. A. Reijers, Fundamentals of Business Process Management, 2018

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	12 hrs	0 hrs	12 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Autumn Semester

# Mathematics V: probabilities and statistics [BDM\_MATHSV]

LEAD PROFESSOR(S): Claire BRECHETEAU

#### Requirements

#### Objectives

Know and understand the basic concepts in probability and statistics. Know how to compute classical probabilities, based on numeration or on the Bayes rule. Master classical probability distributions. Being able to compute statistical descriptors (mean, variance, median, mode, quantiles, cumulative distribution functions). Being able to compute basic confidence intervals, and being able to apply classical statistical tests on the mean.

#### **Course contents**

- 1 Bases on Sets theory and Enumeration
- 2 Introduction to probabilities : from events to random variables
- 3 Discrete and continuous random variables
- 4 Convergence of random variables and Limit theorems
- 5 Estimators in Statistics and Confidence intervals
- 6 Statistical tests

#### Course material

Erwin Kreyszig, Advanced Engineering Mathematics, 10th edition, John Wiley & Sons, 2010

#### Assessment

Collective assessment: EVC 1 (coefficient 0.3)

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	12 hrs	12 hrs	0 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Autumn Semester

# Negotiation [BDM\_NEGO]

LEAD PROFESSOR(S): Luisa ROCHA DA SILVA

#### Requirements

#### Objectives

The purpose of this module is to emphasize the importance of sales for an organization by linking the sales activity to strategic use cases.

#### **Course contents**

Understand the role of sales act as revenue driver for an organization Understand the role of sales ressources as a key element of a development strategy Know about Inbound and Outbound sales Understand sales cycle and customer journey

#### **Course material**

The day after tomorrow: How to survive in times of radical innovation (Peter Hinsen) Challenger sales: Taking Control of the customer conservation (Matthew Dixon & Brent Adamson) Lean Startup (Eric Ries)

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Autumn Semester

# French [BDM\_FLEIII]

LEAD PROFESSOR(S): Caroline WEISSMANN / Silvia ERTL

#### Requirements

#### Objectives

Work : speak about experience after the internship last summer.
Expatriation/live abroad : advantages/disavantages
Life in 10 years : imagine his professionnal and professionnal life in the future.
Project: imagine a firm / an app.

#### **Course contents**

The course activities cover a whole range of practical language and communication exercices that span written and oral comprehension and expression.

#### **Course material**

-Digital tools

-Written and televised press, Internet

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
French	4	0 hrs	32 hrs	0 hrs	0 hrs	0 hrs





2024/2025 - Year 2 - Autumn Semester

# Chinese [BDM\_LV2II]

LEAD PROFESSOR(S): Jianping GUNST

#### Requirements

#### Objectives

This teaching aims to train the learner to be able to hold a conversation in various and real situations in China, for example asking for information at the hotel, at the station, ordering food at the restaurant, etc. Also discover Chinese culture and society.

#### **Course contents**

Phonetics Learn to write characters Manual Sequence 0.0- 1.6

#### **Course material**

Méthode d'Initiation à la Langue et à l'Ecriture chinoises (Joël BELLASSEN)

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	4	0 hrs	32 hrs	0 hrs	0 hrs	0 hrs





2024/2025 - Year 2 - Spring Semester

# Artificial intelligence fundamentals [BDM\_AI]

LEAD PROFESSOR(S): Luisa ROCHA DA SILVA

#### Requirements

#### Objectives

The aim of this course is to offer a general overview of the three blocks upon which this No Code AI and Machine Learning program is focused.

- Understanding the data: What's the data telling us?
- Prediction: What is going to happen?
- Decision-making: What should we do?

#### **Course contents**

There's a great deal of interest in learning and applying Machine Learning and AI techniques across various industries. One thing that is holding people back is the need to learn coding. Even for experienced programmers, implementing Data Science models manually using code can be slow, -frustrating, and error-prone. But now, technological advances have enabled new no-code solutions. The no code approach provides the same key concepts and logical reasoning behind using A[ algorithms to build Data Science solutions without requiring coding.

A few techniques you will be implementing in this program are:

- Clustering Techniques
- Supervised Learning Regression
- Supervised Learning Classification
- Tree-Based Models and Ensemble Techniques
- Natural Language Processing
- Recommendation Systems
- Deep Learning Techniques

#### Course material

1. Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies – By John D. Kelleher, Brian Mac Namee, Aoife D'Arcy

2. Artificial Intelligence: A Modern Approach – ByStuart Russell Peter and Norvig

3. Machine Learning: The New AI – By Ethem Alpaydin

#### Assessment

Individual assessment: EVI 1 (coefficient 1)

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs

This syllabus is not a legally binding document. Content is subject to change year on year.





2024/2025 - Year 2 - Spring Semester

# AI and data mining for digital business [BDM\_AI4BUS]

LEAD PROFESSOR(S): Luisa ROCHA DA SILVA

#### Requirements

#### Objectives

The main objectives of this course are:

- To review key statistical concepts using calculus, specifically focusing on Maximum Likelihood Estimation.
- To understand the fundamentals and core principles of Machine Learning.
- To gain proficiency in a variety of Supervised Learning Algorithms, including Linear Regression, Logistic Regression, and Tree-based models.
- To explore advanced topics in Machine Learning such as Ensemble Learning and Neural Networks.
- To apply Machine Learning techniques in practical scenarios, exemplified by a project on Credit Risk Prediction.

#### **Course contents**

This course provides a comprehensive introduction to Machine Learning, covering both theoretical fundamentals and practical applications. It starts with a statistics review, emphasizing calculus-based approaches like Maximum Likelihood Estimation. The course then delves into the core areas of Machine Learning, including both foundational topics and advanced techniques such as Ensemble Learning and Neural Networks. A significant portion of the course focuses on practical implementation, with an emphasis on applying learned concepts to real-world problems like Credit Risk Prediction. The course is structured to build from basic to complex concepts, preparing students for both academic advancement and professional application in the field of Machine Learning.

#### Course material

- 1. Machine Learning: a Probabilistic Perspective (Kevin Murphy)
- 2. Pattern Recognition and Machine Learning (Christopher Bishop)
- 3. Deep Learning (Ian Goodfellow Yoshua Bengio Aaron Courville)

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Spring Semester

# Data cleansing and data quality [BDM\_CLEAN]

LEAD PROFESSOR(S): Luisa ROCHA DA SILVA

#### Requirements

#### Objectives

The objective of this course is to provide students with the basics of data cleaning by relying on a variety of practices in Python on different data formats. The skills acquired should facilitate the manipulation and cleaning of dirty data in a professional environment while understanding the ins and outs of data quality issues.

#### Course contents

The data cleaning course is an exploration of data quality issues across over twenty possible data formats. The course also covers the reasons behind poor quality as well as data cleaning techniques.

#### **Course material**

1. https://www.oreilly.com/library/view/python-data-cleaning/9781800565661/

2. https://www.oreilly.com/library/view/practical-python-data/9781492091493/

3.https://pandas.pydata.org

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Spring Semester

# Cloud computing [BDM\_CLOUD]

LEAD PROFESSOR(S): Hugues DIGONNET

#### Requirements

#### Objectives

The goal of this course is to learn how to create and deploy Cloud applications.

#### Course contents

This course includes an introduction to Linux, containerization, and parallelism, followed by practical implementation in the cloud: definition and types of cloud (private, public, hybrid), cloud service models (SaaS, PaaS, IaaS), risks and opportunities of the cloud (security management, instant access to computing and storage resources – elasticity), creation and deployment of applications (Google Cloud, Amazon, Azure, etc.).

#### **Course material**

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Spring Semester

# Data analysis [BDM\_DATAANAL]

LEAD PROFESSOR(S): Luisa ROCHA DA SILVA

#### Requirements

#### Objectives

- Understand the principles of data analysis.
- Learn various methods for data visualization and exploration.
- Apply statistical techniques to interpret data.
- Apply machine learning to prepare and explore data.

#### **Course contents**

- Chapter 1: Introduction to Data Analysis
- What is data analysis?
- Importance of data-driven decision making.
- Overview of data analysis process.

Chapter 2: Data cleaning and Preparation

- Handling Missing Data: Imputation, deletion, or estimation.
- Handling Imbalanced Data: oversampling, undersampling, synthetic data.
- Outlier Detection and Treatment: Identify and correct anomalies.
- Data Transformation: Normalization, standardization, or scaling.
- Feature Engineering: Create new variables to enhance analysis.

Chapter 3: Exploratory Data Analysis (EDA)

- Univariate Analysis: Study of individual variables.
- Bivariate Analysis: Relationships between pairs of variables.
- Multivariate Analysis: Simultaneous study of multiple variables.
- Segmentation and Clustering: Grouping data to uncover patterns.

#### Chapter 4: Dimensionality Reduction

- Principal Component Analysis (PCA): Reduces dimensionality while preserving variance
- Feature Selection: filter, wrapper, embedded methods
- Visualization and correlation analysis

#### Practical sessions:

- Dealing with missing data and imbalanced data
- Data transformation and data cleaning
- Regression analysis
- Clustering analysis
- Dimensionality reduction

#### Course material



1. "Exploratory Data Analysis" by John W. Tukey

2. "Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking" by Foster Provost and Tom Fawcett

3. "Python for Data Analysis", by Wes McKinney

#### Assessment

Individual assessment: EVI 1 (coefficient 1)

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs

This syllabus is not a legally binding document. Content is subject to change year on year.





2024/2025 - Year 2 - Spring Semester

# Data mining [BDM\_DATAMIN]

LEAD PROFESSOR(S): John MONEER ISHAK HANNA

#### Requirements

#### Objectives

The aims of this course are to presents to the students well-known data mining techniques and their application areas. Specifically, the course demonstrates basic concepts, principals and methods of data mining. It also demonstrates the process of Knowledge Discovery in Databases (KDD) and presents a review of available tools.

#### **Course contents**

The course covers various topics essential to data analysis and implementation using Python, focusing on real-life examples. It begins with an exploration of data fundamentals and practical Python applications. The classification section starts with basic concepts and progresses to specific algorithms, demonstrated through real-world examples using both Weka and Python. Basic concepts of association analysis and cluster analysis are also included. Additionally, the course may also covers anomaly detection techniques.

#### Course material

1. Han, J., Kamber, M., Pei, J. Data Mining: Concepts and Techniques, Third Edition. – Morgan Kaufmann Publishers, 2011. – 740 pp

2. Johannes Ledolter. Data Mining and Business Analytics with R. Wiley. 2013

3. Hall, M., Witten, Ian H., Frank, E. Data Mining: practical machine learning tools and techniques. – 2011. – 664 pp.

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Spring Semester

# Machine and deep learning [BDM\_MLDEEPL]

LEAD PROFESSOR(S): Hachem MADMOUN

#### Requirements

#### Objectives

• Understand the Fundamentals:

Comprehend the basic concepts and techniques of machine learning and deep learning.

0 hrs

Explore various types of machine learning algorithms including supervised, unsupervised, and self supervised learning.

Develop Practical Skills:

Gain hands-on experience with key machine learning libraries and frameworks such as scikit-learn, TensorFlow, and PyTorch. Implement and evaluate machine learning models on real-world datasets

• Evaluate Model Performance:

Learn to assess and improve the performance of machine learning models using various metrics and validation techniques. Explore techniques for model optimization and regularization.

#### **Course contents**

This course provides a comprehensive introduction to machine learning and deep learning, covering both theoretical foundations and practical implementations. Students will learn to design, build, and evaluate machine learning models, as well as explore advanced deep learning techniques. The course includes lectures, hands-on lab sessions, and projects that apply these techniques to solve real-world problems. By the end of the course, students will be equipped with the knowledge and skills to pursue further studies or careers in machine learning and deep learning.

#### **Course material**

#### Assessment

English

Individual assessm	Individual assessment: EVI 1 (coefficient 1)							
LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM		

24 hrs

0 hrs

0 hrs

2 hrs





2024/2025 - Year 2 - Spring Semester

# Natural Language Processing [BDM\_NLP]

LEAD PROFESSOR(S): Luisa ROCHA DA SILVA

#### Requirements

#### Objectives

- Acquire the fundamental concepts of NLP.
- Understand the various applications and tasks related to NLP.
- Acquire classical and advanced techniques of NLP.
- Be able to develop NLP algorithms using Python and PyTorch.

#### **Course contents**

Fundamentals and evolution of NLP. Preprocessing of textual data. Classic NLP tasks. Statistical techniques for NLP. Topic modeling. Machine learning for NLP. Distributed representations. Deep learning for NLP (RNN, LSTM, seq2seq). Self-supervised learning and Transformers. Multimodal learning.

#### **Course material**

- 1. Foundations of statistical natural language processing. C. Manning, H. Schutze. MIT Press, 1999.
- 2. Natural Language Processing with Python. S. Bird, E. Klein. O'Reilly, 2009.
- 3. Natural Language Processing with Transformers. L. von Werra, T. Wolf. O'Reilly, 2022.

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Spring Semester

# Scientific visualization [BDM\_SCIVIZ]

LEAD PROFESSOR(S): Lucas LESTANDI

#### Requirements

#### Objectives

At the end of the course, students should have basic skills in the two main aspects of data visualization.

1) They will be able analyze and explore data visualization tools.

They will know how to explore data with multiple tools. Among the most widespread, they will have know when and how to use

- a spreadsheet software (excel, Tableau)
- Scripted plotter in python (matplotlib)
- Scientific visualization software (paraview) both in GUI mode and scripted mode
- or decide to go beyond with specialized tools like R or dashboard creators (Graphana)

2) They will be able to process and display efficiently data of multiple kinds including:

- time series
- categorical data
- spatial & geographical data
- high dimensional data
- scientific data

3) Students will acquire essential abilities and effective methodologies for sharing the conclusions of the analysis with external parties through high quality graphical communication.

#### **Course contents**

- 1) Introduction
  - General context with examples, visuals
  - Review of the various kinds of graphs and visualisations
  - Visualization tools categories
  - Segmenting the problem : from data collection (exploratory data analysis) to sharing
- 2) High level data visualization : the alpha and omega of dataviz
  - The old world : Spreadsheets (excel) and Powerpoint
  - Modern tools : dashboards like Tableau, graphana (and BI tools)
- 3) Scientific visualisation : Exploring data visually
  - Visualizing "scientific data" (with python)
  - Statistical
  - (very) high dimensional data (many-to-many)
  - Geographical
  - Viewing 3D scientific data. How it works : from points and fields to display

4) Visualization for humans



- How we perceive things
- Conveying meaning
- How to display data efficiently
- Data storytelling

#### **Course material**

Storytelling with data, Cole Nussbaumer Knafflic Scientific Visualization, Nicolas Rougier Data Points, Visualization means something, Nathan Yau

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs





2024/2025 - Year 2 - Spring Semester

# Systems security information [BDM\_SSI]

LEAD PROFESSOR(S): Pablo BONDIA-LUTTIAU

#### Requirements

#### Objectives

Know the main threats linked to big data Know how to effectively protect data in terms of confidentiality and integrity Know the main security mechanisms implemented on the Internet Know how to destroy and recover data

#### **Course contents**

- Introduction to the main security principles
- Cryptography
- Passwords and hashes
- SQL and Databases
- IT
- Scans and exploitation
- Final TP

#### Course material

ANSSI : https://cyber.gouv.fr/en

#### Assessment

LANGUAGE OF	ECTS CREDITS	LECTURES	TUTORIALS	LAB	PROJECT	EXAM
English	3	0 hrs	24 hrs	0 hrs	0 hrs	2 hrs