Master Data Sciences For Societal Challenges

Antenne de Blois

Des experts en ingénierie et sciences des données

université de TOURS | Faculté des sciences & techniques

www.info.univ-tours.fr/Blois

master-info-blois@univ-tours.fr

Université de Tours
Faculté de Sciences et Techniques
Département d'Informatique
3 place Jean Jaurès
41000 Blois

Curriculum Description of CS Master on Data Sciences for Societal Challenges

MASTER Sciences, Technologies, Santé

MENTION Informatique

PARCOURS Data Sciences for Societal Challenges

In 2009, Turing award winner Jim Gray spoke of data science as a fourth paradigm of science (empirical, theoretical, computational and data-driven) arising from and capitalizing on the huge amount of data that is now available for investigation. The confluence of the availability of data and increasingly sophisticated tools, processes, and algorithms for analyzing and drawing knowledge and insight from data has impacted every area of scientific engagement. It has also opened exciting new opportunities for interdisciplinary work across the many fields including (but certainly not limited to) computer science, mathematics, statistics, and information science from which it draws foundational knowledge.

as read in
https://www.acm.org/binaries/content/assets/education/curricula-recommendations/dstf_ccdsc2021.pdf

Practical information

Diploma training

Accessible for initial, apprenticeship and continuous training

Duration: 2 years

Internship: Optional in 1st year (up to 4 months), mandatory in 2nd year (6 months)

Teaching languages: English and French

Training location: Blois

Responsible: Thomas Devogele (thomas.devogele@univ-tours.fr)

Contents

Semester 7

UE 7.1 Statistics for Data Sciences

EP 7.1.1 Statistics

Prerequisites: Basics in linear algebra, matrix calculations, functions of several variables, random variables and basic statistics (1st and 2nd order statistics such as mean and variance), and correlation and linear regression.

Hours: lectures: 9, tutorials: 13, labs: 6

ECTS: 4

Objectives: The objectives of this course are to present the main methods and techniques to analyze univariate and multivariate datasets. Noticeably, a focus is made on variance analysis, statistical tests and correlation measures.

Acquired Skills and Knowledge:

- Analyzing a problem and identifying features properties based on their distributions, or their correlation
- Identifying the main distribution families (uniform, normal, exponential)
- Implementing test on proportion (Bernoulli), tests on the average (bilateral, unilateral), on variance (two-sided, one-sided), two-sample comparison test, ANOVA test
- Assessing the independence and the significance of results based on statistical tests: Chi-square test (adequacy to a discrete law, independence between two samples), test of adequacy to the normal law

UE 7.2 Data Management

EP 7.2.1 DB Modeling and Querying Essentials

Prerequisites: This course assumes a basic knowledge of relational database querying and design: relational model, conjunctive queries, and SQL.

Hours: lectures: 7.5, tutorials: 6, labs: 7.5

ECTS: 3

Objectives: Study the logical properties of relational database languages. Understand the various classes of queries along with their complexity and properties. Understand the underlying structures and mechanisms used by database engines (indexes, data organization, transaction)

Acquired Skills and Knowledge:

- Advanced querying (extend relational queries, views, recursive queries)

- Understanding of query classes and their complexity
- Understanding of transactions and their properties (ACID)
- Basic understanding of index structures (trees, hashing)
- Basic knowledge of database normalization principles (3NF, etc.)

EP 7.2.2 Data Warehouse Modeling

Prerequisites: This course assumes a solid knowledge of relational database querying and design: relational model, SQL, entity-relationship model, normalization

Hours: lectures: 10,5, tutorials: 6, labs: 7,5

ECTS: 3

Objectives: Study database modeling technologies for business intelligence. Understand architecture, models and life cycle of a data warehousing system. Experience the design of a small data warehouse with current tools.

Acquired Skills and Knowledge:

- Understanding of business intelligence modeling and querying needs, in particular, the OLAP paradigm
- Understanding of data warehousing architectures and design methodologies
- Understanding of multidimensional models
- Mastering of conceptual and logical design techniques
- Practical design experience with current tools.

EP 7.2.3 Data Visualization

Prerequisites – Notions on ergonomics of digital technologies

Hours: lectures: 3, tutorials: 6

ECTS: 2

Objectives: Knowing how to represent efficiently massive or aggregate data for a given communication aim, by choosing an appropriate representation system and an effective practical implementation.

Acquired Skills and Knowledge:

- Knowing how to choose a representation system (type of chart) adapted to any specific communication aim related to the visualization of massive data.
- Knowing how to implement practically a given chart in all its complexity (colours, chart orientation, data ordering, data scales, etc.)
- Being familiar with the main ergonomic guidelines for data visualization and knowing how to apply them practically.

Content:

- Introduction: why data visualization is important, and its relation with communication aims.
- Types of representation systems and their relationships with data types and communication aims : scatter plots, pie charts, bar charts, density plots, box and whiskler plots, index charts, layer graphs, horizon graphs, slope charts, tree maps, sunbursts, relational graphs.
- Ergonomic guidelines : perceptive and cognitive laws and their application to data visualization.

UE 7.3 Data Mining

EP 7.3.1 Data Preparation

Prerequisites: This course assumes a basic knowledge of relational database querying and design (relational model, SQL, constraints), and programming skills.

Hours: lectures: 3, tutorials: 1.5, labs: 19.5

ECTS: 4

Objectives: Study the main data preparation techniques used in data management and data mining projects. Experience the development of data pipelines with current tools.

Acquired Skills and Knowledge:

- Understanding of data transformation needs and data quality issues.
- Understanding of data profiling and data analysis techniques (distributions, correlations...).
- Understanding of data extraction, transformation and loading (ETL) environments.
- Understanding of feature selection and engineering techniques.
- Design and implementation of data pipelines, including data cleaning, data aggregation and computation of indicators.
- Practical design experience with current tools.

EP 7.3.2 Data Mining 1

Prerequisites: Statistics, Linear Algebra, Matrix calculus, Functions of several variables, Algorithmic

Hours: lectures: 12, tutorials: 9, labs: 10.5

ECTS: 4

Objectives: Know the main steps and tasks of a knowledge discovery process. Understand the different types of supervised and unsupervised learning problems. Know how to implement some common clustering, classification and pattern-finding techniques. Set up an evaluation protocol for a data mining task.

Acquired Skills and Knowledge:

- Machine Learning Problem and ML evaluation
- Simple classifiers: Neural Networks, Decision Trees, SVM
- Clustering algorithms: k-means, hierarchical clustering, DB-SCAN
- Pattern Mining: A priori, FP-Growth

UE 7.4 Data Science Project

EP 7.4.1 Scientific Method

Prerequisites: This course assumes a basic knowledge of computer science, communication and English.

Hours: lectures: 4.5

ECTS: 1

Objectives: Learn to synthesize, organize, and present scientific and technical information related to the data science domain, to make it clear to colleagues, and discuss it objectively

Acquired Skills and Knowledge:

- Good understanding of the state of the art in a specific subdomain of data science
- Synthesis, organization, and presentation of scientific and technical information
- Practical experience during a Data Science project (see EP 7.4.2)

EP 7.4.2 Data Science Project 1

Prerequisites: This course assumes a solid knowledge of relational database querying and design, as well as programing, and basic knowledge of project management.

Hours: tutorials: 36. Students have additional dedicated time and room to work on this project.

Remark: For initial training students, the required involvement is higher, and students have two additional days per week to work on this project.

ECTS: 5

Objectives: Summarize the knowledge acquired during the semester through a concrete large-scale project in the field of data sciences. Concretely, the project includes the study of a data-oriented problem (with business requirements involving both statistical analysis and machine learning), preparation and integration of datasets, modeling and implementation of a data warehouse architecture.

Acquired Skills and Knowledge:

- Reinforcement and application of skills of EP 7.2.2, 7.1.1, 7.3.1 and 7.4.1
- Project management and planning
- Teamwork
- Written and oral presentation of project advancement and deliverables, including demos

UE 7.5 Transverse Skills

EP 7.5.1 Seminars 1

Prerequisites: This course assumes students are following (and have prerequisites) for the other courses of the semester

Hours: Approximately 12 hours of tutorials (depending on invited speakers)

ECTS: 1

Objectives: Provide an overview of trends in the data science field in terms of research and innovation. Present application contexts and real use cases of the theoretical methods seen in class.

Seminars are a series of invited talks and lectures by renowned professors and researchers from the database and machine learning communities, and professionals from varied industrial sectors.

Acquired Skills and Knowledge:

- Reinforcement of skills of UE 7.1, 7.2 and 7.3.

EP 7.5.2 Professional Communication 1

Prerequisites: B1 to B2 English level (at least 600 Toeic score, or equivalent)

Hours: tutorials: 24

ECTS: 3

Objectives: Gain confidence in speaking and be clear in writing

Acquired Skills and Knowledge:

- Be able to present orally in front of your client,
- Be able to write professional documents

Semester 8

UE 8.1 Foundations of Data Sciences

EP 8.1.1 Operation Research

Prerequisites: basics in linear algebra, algorithmics, basic data structures, complexity of algorithms

Hours: lectures: 12, labs: 6

ECTS: 2

Objectives: Understanding mathematical programming and graph theory tools to model and solve problems.

Acquired Skills and Knowledge:

- Analyze the complexity and the difficulty of a problem
- Identify most appropriate algorithms to solve a problem
- Solve real problems using mathematical programming as well as heuristics

EP 8.1.2 Advances Statistics

Prerequisites: Statistics, linear algebra

Hours: lectures: 7, tutorials: 11, labs: 3

ECTS: 3

Objectives: The main objective is to understand multivariate analysis methods such as Principal Component Analysis, Factor Correspondence Analysis and Multiple Correspondence Analysis. Apart from principles of PCA, the course details the main steps to leverage PCA such as matrix calculation, proof, indicator and interpretation.

Acquired Skills and Knowledge:

- Analyzing and interpreting numerical data based on variance (PCA)
- Analyzing and interpreting categorical data (FCA)
- Displaying or summarizing a set of data in two-dimensional graphical form

UE 8.2 Data Knowledge and Management

EP 8.2.1 Web Data and Knowledge Representation

Prerequisites: basics in Web, databases, data representation, logics

Hours: lectures: 9, tutorials: 10,5, labs: 9

ECTS: 4

Objectives: Know the principles of data and knowledge engineering. Know the Web standards for Web data and knowledge Representation

Acquired Skills and Knowledge:

- Semi-structured Web data: schemas and integrity constraints
- Web communities and standard XML language proposals
- From schemas to ontology: knowledge engineering, ontology design
- Description logics
- Semantic Web standards for ontologies and knowledge graphs
- Web communities and ontologies

EP 8.2.2 Internals of Databases

Prerequisites: EP 7.2.1

Hours: lectures: 10.5, tutorials: 9, labs: 7.5

ECTS: 4

Objectives: Study the fundamentals of the implementation of a relational database management system. Understand the principles of query evaluation, implement a query optimization strategy and administer a relational database.

Acquired Skills:

- Understanding of the query evaluation pipeline
- Understand an influence the behavior of the query optimizer
- Configure and optimize a DBMS for a particular workload

UE 8.3 Advanced Data Mining

EP 8.3.1 Data Mining 2

Prerequisites: Statistics, Linear Algebra, Matrix calculus, Functions of several variables, Algorithmic

Hours: lectures: 15, tutorials: 4.5, labs: 15

ECTS: 4

Objectives: Understand the foundations of ML models and empirical risk minimization (Vapnik Chervonenkis). Understand fundamentals of neural networks with forward and backward pass and gradient descent principle. Understand the adaptation of clustering to big data with kernel k-means, parallel k-means or stream-clustering algorithms. Know how to leverage sequential pattern-mining techniques.

Acquired Skills and Knowledge:

- Understanding the fundamentals of ML and the learnability problem
- Understanding big data clustering problems and techniques
- Leveraging techniques to deal with complex data such as graphs or sequences

EP 8.3.2 Data Quality

Prerequisites: This course assumes a solid knowledge of (structured and semi-structured) data representation models (especially concepts seen in EP 7.2.1 and 7.2.2) and basic knowledge on data preparation issues (studied in EP 7.3.1). Preliminary knowledge on statistical and data mining techniques (studied in EP 7.1.1. and 7.3.2) is also useful.

Hours: lectures: 9, tutorials: 1.5, labs: 10.5

ECTS: 2

Objectives: Study the main concepts and techniques for assessing and improving data quality, by deeply studying key quality dimensions, methodologies, algorithms and specialized tools

Acquired Skills and Knowledge:

- Understanding of data quality terminology and methodologies.
- Understanding, identification and usage of main quality dimensions and metrics.
- Understanding of major techniques for data quality assessment and improvement.
- Understanding of major techniques for data integration and fusion.
- Practical experience for assessing and improving data quality using current tools.

UE 8.4 Data Science Project

EP 8.4.1 Data science project 2

Prerequisites: This course assumes a solid knowledge of topics of data science project (UE 7.4) and data mining (EP 7.3.2)

Hours: tutorials: 36. As in EP 7.4.1, students have dedicated time and room to work on the project.

ECTS: 5

Objectives: Summarize the knowledge acquired during both semesters through a concrete large-scale project in the field of data sciences. Concretely, the project continues the data science project of previous semester (EU 7.4) and includes feature engineering, implementation and tuning of data mining models, and evaluation of model performance.

Acquired Skills and Knowledge:

- Reinforcement and application of skills of EP 7.3.2, 8.3.1, 8.1.2 and 8.3.2
- Project management and planning
- Teamwork
- Written and oral presentation of project advancement and deliverables, including demos

UE 8.5 Transverse Skills

EP 8.5.1 Seminars 2

Prerequisites: This course assumes students are following (and have prerequisites) for the other courses of the semester

Hours: Approximately 12 hours of tutorials (depending on invited speakers)

ECTS: 1

Objectives: Provide an overview of trends in the data science field in terms of research and innovation. Present application contexts and real use cases of the theoretical methods seen in class.

Seminars are a series of invited talks and lectures by renowned professors and researchers from the database and machine learning communities, and professionals from varied industrial sectors.

Acquired Skills and Knowledge:

- Reinforcement of skills of UE 8.1, 8.2 and 8.3.

EP 8.5.2 Professional Communication 2

Prerequisites: B1 to B2 English level (at least 600 Toeic score, or equivalent)

Hours: tutorials: 24

ECTS: 3

Objectives: Apply appropriate communication skills across settings, purposes, and audiences.

Acquired Skills and Knowledge:

- Use technology to communicate effectively in various settings and contexts.
- Demonstrate appropriate and professional ethical behavior.
- Demonstrate critical and innovative thinking.
- Display competence in oral, written, and visual communication.
- Respond effectively to cultural communication differences.
- Communicate ethically.
- Demonstrate positive group communication exchanges.

UE 8.6 Data Science Applications

EP 8.6.1 Intership or Report

Prerequisites: This course assumes students are following (and have prerequisites) for the other courses of the semester, in particular those related to the chosen internship or report subject.

Hours: tutorials: 20. Students should provide at least 4 weeks of work in their internship or report.

ECTS: 2

Objectives: Put in practice the knowledge acquired during the semester to deepen the study a specific subject, either as an internship or by writing a project report.

Acquired Skills and Knowledge:

- Reinforcement of skills of UE 8.1, 8.2 and 8.3.
- Application of the knowledge acquired during the semester
- Write a summary report

Semester 9

UE 9.1 Data Variety

EP 9.1.1 Spatio-temporal Analysis

Prerequisites: database, statistic, data mining

Hours: lectures: 13.5, tutorials: 7.5, labs: 9

Objectives: Understanding the notions time series, geographic information systems (GIS), analysis temporal or spatial data

Acquired Skills and Knowledge:

- Time series decomposition
- Time series prediction (ARMA, ARIMA, LSTM)
- Spatial analysis
- Time and spatial data clustering

EP 9.1.2 Stochastic Processes

Prerequisites: basics in probability (random variables, probability distribution)

Hours: lectures: 9, tutorials: 3, labs: 1,5

Objectives: Understanding the notions of expectations, joint probability distributions, conditional probabilities and building upon these notions the concepts of Markov Decision Processes.

Acquired Skills and Knowledge:

- Formalizing the notions of expectations, joint /conditional probability
- Understanding and implementing simple Markov Decision Processes

UE 9.2 Advanced Data Management

EP 9.2.1 Warehousing, Data Lakes and Polystores

Prerequisites: EP 7.2.1 and EP 7.2.2

Hours: lectures: 10,5, tutorials: 4,5, labs: 9

Objectives: Discover the different types of 'NoSQL' databases and their application to multimodal data storage. Understand basic properties of distributed systems (e.g. Eventual consistency)

Acquired Skills and Knowledge:

- Introduction to data lakes and 'NoSQL' database systems
- Loading data into and querying document-oriented databases (MongoDB), graph-oriented databases (Neo4J) and Key/Value stores.
- Introduction to polystores (Apache Drill)
- Logical properties of distributed systems

EP 9.2.2 Web Data and Knowledge Exploitation and Management

Prerequisites: EP 8.2.1. Web Data and Knowledge Representation

Hours: lectures: 10,5, tutorials: 10,5, labs: 12

Objectives: Know the principles of - knowledge - graph querying and management. Know what is the Linked Open Data cloud and how to exploit the main public knowledge resources.

Acquired Skills and Knowledge:

- Querying of knowledge graphs
- LOD principles and LOD's resources : wikidata, dbpedia, others
- Knowledge graphs management: triplestores
- Property graphs: property graph databases, property graph querying

UE 9.3 Deep Learning

EP 9.3.1 Deep Learning

Prerequisites: Machine Learning fundamentals, introduction to neural networks in 8.3.1, knowledge of calculus, linear algebra, statistics, and probabilities is beneficial, along with basic Python experience

Hours: lectures: 7.5, labs: 10.5

Objectives: Gaining a practical understanding of the internals of deep neural architectures such as convolutional networks, recurrent networks, attention mechanisms and transformers.

Acquired Skills and Knowledge:

- Knowing how to use TensorFlow and Pytorch Python libraries
- Understanding the fundamentals of representation learning
- Illustrating convolutional neural networks on simple image analysis

EP 9.3.2 Hackathon

Prerequisites: EP 7.1.1 / EP 7.3.1 / EP 7.3.2 / EP 8.1.2 / EP 8.2.1 / EP 8.3.1 / EP 8.3.2

Hours: A full day

Objectives: The Hackathon is a full day dedicated to applying machine learning approaches discovered by the students during their master's degree. The students arrive in the morning and discover a ML oriented task and dataset. They form teams and start data cleaning, feature engineering and ultimately complete the ML task. The data and task may be linked to active research projects from the lab or partner institutions and companies.

Acquired Skills and Knowledge:

- Teamwork
- Quick iteration and turnaround on a proof of concept
- Reinforce and apply skills of EP 7.1.1 / EP 7.3.1 / EP 7.3.2 / EP 8.1.2 / EP 8.2.1 / EP 8.3.1 / EP 8.3.2

UE 9.4 Data Science Project

EP 9.4.1 Data Science Project 3

Prerequisites: This course assumes a solid knowledge of topics of data science projects (UE 7.4 and 8.4). Several subjects are proposed, each one having additional prerequisites involving topics learned during first year courses.

Hours: tutorial: 40. Students have two days a week dedicated to the project.

Objectives: Manage a full project related to data science topics from the specification of the needs to the delivery of a completely functional software solution.

Acquired Skills and Knowledge:

- Mastering of data science concepts, languages, frameworks and API to solve the problem at hand.
- Gain in experience by mobilizing transverse skills acquired during all master courses
- Personal project management and planning
- Written and oral presentation of project advancement and deliverables, including demos

UE 9.5 Transverse Skills

EP 9.5.1 Seminars 3

Prerequisites: This course assumes students are following (and have prerequisites) for the other courses of the semester

Hours: Approximately 12 hours of lectures (depending on invited speakers)

Objectives: Provide an overview of trends in the data science field in terms of research and innovation. Present application contexts and real use cases of the theoretical methods seen in class.

Seminars are a series of invited talks and lectures by renowned professors and researchers from the database and machine learning communities, and professionals from varied industrial sectors.

Acquired Skills and Knowledge:

- Reinforcement of skills of UE 9.1, 9.2 and 9.3.

EP 9.5.2 Professional communication 3

Prerequisites: B1 to B2 English level (at least 750 Toeic score, or equivalent)

Hours: tutorials: 24

Objectives: Apply appropriate communication skills across settings, purposes, and audiences.

Acquired Skills and Knowledge:

- Use technology to communicate effectively in various settings and contexts.
- Demonstrate appropriate and professional ethical behavior.
- Demonstrate critical and innovative thinking.
- Display competence in oral, written, and visual communication.
- Respond effectively to cultural communication differences.
- Communicate ethically.
- Demonstrate positive group communication exchanges.

Semester 10

UE 10.1 Data Variety

EP 10.1.1 Natural Language Processing (NLP)

Prerequisite: Machine Learning (MSc.), Deep learning (MSc.), Formal languages (BSc.)

Learning outcomes:

- Understand the nature of natural language data and their modeling,
- Manipulate some existing NLP tools with a deep understanding of the underlying processes involved,
- Develop sample NLP applications, potentially through the integration of existing tools, in order to process textual data.

Content:

- Introduction: NLP applications in the digital society,
- Nature of language data, foundations of linguistic modeling,
- Key processing techniques for textual data : tokenizsation, stemming, lemmatization, POS tagging, named entity recognition, syntactic parsing
- Encoding textual data and format manipulation,
- State-of-the-art NLP techniques heavily used in applications: symbolic, machine learning and deep learning approaches. Comparative study of these approaches along several criterion: accuracy (robust AI), transparency and explicativity (understandable AI), portability and domain independence (adaptable AI), computational cost (sustainable AI).

UE 10.2 Advanced Data Management

EP 10.2.1 Large Scale Data Analysis

Prerequisites: Java, Database, Data Mining

Hours: lectures: 12, labs: 12

Objectives: Understanding the platforms and methods for large scale data analysis.

Acquired Skills and Knowledge:

- The Hadoop platform and the MapReduce programming model
- Understanding and implementing MapReduce jobs for different data analysis tasks
    - Basic statistics
    - Pattern mining
    - Skyline retrieval

EP 10.2.2 Knowledge Quality

Prerequisites: EP 8.2.1. Web Data and Knowledge Representation and EP 9.2.2 Web Data and Knowledge Exploitation and Management

Hours: lectures: 6, tutorials: 3, labs: 6

Objectives: Know how to represent quality as knowledge. Know how to generate qualitative knowledge graphs. Know how to evaluate knowledge graph's quality.

Acquired Skills and Knowledge:

- Knowledge graph's Quality Dimensions
- Ontologies for expressing datasets quality
- FAIR principles

- Tools for evaluating and improving KG quality

UE 10.3 User Centric Approaches

EP 10.3.1 Ranking and Recommendation

Prerequisites: Machine Learning, Data Mining, Algorithmic Knowledge

Hours: lectures: 10.5, tutorials: 4.5, labs: 10.5

Objectives: Understanding the ranking and recommendation problems in different application domains. Learning existing algorithms and tools. Experience the real data.

Acquired Skills and Knowledge:

- Key notions of ranking and recommendation
- Collaborative recommendation techniques
- Skyline and Top-k retrievals

EP 10.3.2 Exploratory Data Analysis

Prerequisites: This course assumes a solid knowledge on data visualization (studied in EP 7.2.3) and data mining (studied in EP 7.3.2. and 8.3.1), as well as basic knowledge of data management (studied in EP 7.2.1, 7.2.2 and 8.2.2) and knowledge management (studied in EP 8.2.1, and 9.2.2).

Hours: lectures: 10.5, labs: 13.5

Objectives: Study the whole data life cycle, from data collection and analysis to data visualization, narrating data stories. Use of advanced techniques for discovering, explaining, structuring and visualizing insights. Experience the narration of a data story.

Acquired Skills and Knowledge:

- Understanding of the data narration process.
- Understanding of techniques for exploratory data analysis, insight discovery, and interest computation.
- Understanding of explainability techniques.
- Understanding of semi-supervised approaches.
- Understanding of advanced visualization techniques (multidimensional, hierarchical and temporal data).
- Practical experience for crafting data narratives using current tools.

UE 10.4 Data Science Project

EP 10.4.1 Data Science Project 4

Prerequisites: This course assumes a solid knowledge of topics of data science projects (UE 7.4, 8.4 and 9.4), as well as the additional prerequisites of the actual project topic started at UE 9.4.

Hours: tutorial: 40. Students have two days a week dedicated to the project.

Objectives: Manage a full project related to data science topics from the specification of the needs to the delivery of a completely functional software solution. Concretely, the project continues the data science project of previous semester (EU 9.4).

Acquired Skills and Knowledge:

- Mastering of data science concepts, languages, frameworks and API to solve the problem at hand.
- Gain in experience by mobilizing transverse skills acquired during all master courses
- Personal project management and planning
- Written and oral presentation of project advancement and deliverables, including demos

UE 10.5 Transverse Skills

EP 10.5.1 Seminars 4

Prerequisites: This course assumes students are following (and have prerequisites) for the other courses of the semester

Hours: Approximately 12 hours of lectures (depending on invited speakers)

Objectives: Provide an overview of trends in the data science field in terms of research and innovation. Present application contexts and real use cases of the theoretical methods seen in class.

Seminars are a series of invited talks and lectures by renowned professors and researchers from the database and machine learning communities, and professionals from varied industrial sectors.

Acquired Skills and Knowledge:

- Reinforcement of skills of UE 10.1, 10.2 and 10.3.

EP 10.5.2 Ethics

Prerequisite: None

Learning outcomes:

- Evaluate and discuss the impact of a scientific or an engineering work
- Identify the code of ethics (or the ethically related regulation) employed by professional bodies

Content:

- Philosophical foundation of ethics and current ethical approaches : ethics of virtues, deontologic and consequentialist ethics
- Ethics and regulation : soft law vs. hard law, ethical grounding of law (universal principles of human rights)
- Deontolological ethics and computer science: ethics principles (privacy, property, accuracy, fairness, equiy, transparency, accountability, accessibility) and ethical guidelines.
- Consequentialism and computer science: ethics and technological risk (individual, societal and environmental impact of digital technologies.
- Ethics, computer science and data science: case studies. Methodology for a ethical impact assessment of technology
- Law aspect (General Data Protection Regulation, IA act and other reglementations)

UE 10.6 Master Thesis

EP 10.6.1 Master Thesis

EP 10.6.2 Data Science Project 5 (only for apprentices)