



UNIVERSITÀ DI PISA  
Dipartimento di Informatica

Master degree programme in  
**Data Science and Business Informatics**  
(2 years, 120 ECTS)

**Study plan rules (“Regolamento”) and students’ guide**  
Academic Year 2026/27

*Contact for information*

<mailto:datascience@di.unipi.it>

*Web site*

<https://www.di.unipi.it/en/education/mds>

Last update: **July 2, 2027**



---

## Preamble

The two-year Master degree program *Data Science and Business Informatics* is designed to prepare graduates both to master information technologies and to understand the needs of organizations with a specific training in *Business Intelligence* and *Data Science* for decision support.

Most of the courses of the Data Science and Business Informatics degree are taught in English. A small number of courses from the Business Economics and Business Law area are still taught in *Italian* (marked with an asterisk in this document), but it is entirely possible to have the whole study plan built up with courses taught in English, making the course feasible to non Italian-speaking students. Indeed, Italian students must have a sufficient knowledge of English at enrollment time, at least at B2 level.

The master programme requires a solid background, high motivation, and hard working attitude. Concept abstraction, problem solving, formal modeling, mathematical reasoning, basic concepts of computer programming and databases, and a solid grasp of the fundamental aspects of the functioning of business entities are essential characteristics that you should eventually possess. Students shall not underestimate this advice. However, the course is also structured to allow access from students with a variety of backgrounds: computer science, economics / management, and the hard sciences in general (mainly mathematics, physics, statistics). Each student receives a bespoke study plan whereby deficiencies in fundamental areas (say, computer science for students with economic background or economics for students with computer science background) are recovered by means of some specifically designed courses.

Courses assessments usually consist of a written and an oral exam. In the written exam, the student must demonstrate the use of knowledge of the course contents to solve problems. During the oral exam the student must be able to demonstrate knowledge of the course contents and be able to discuss the topics thoughtfully and with propriety of expression. A number of courses, however, substitute the written exam (totally or partly) with project-based and/or laboratory work.

Attendance of courses is not mandatory. Most courses offer support to non-attending students under various forms, e.g., extensive material (lectures, books, slides, software) and/or registration of lectures. However, each lecturer freely decides the level of support that (s)he offers, and succeeding in some courses—in particular, those with higher hands-on components and requiring more laboratory work—may be harder without attending. Furthermore, the course is challenging and its difficulty should not be underestimated. Since part-time students experience lower success rates in exams and a longer time to graduate, we highly recommend that students regularly attend lectures and complete the courses each semester.

Our graduates are highly sought after in the job market. Not only statistics show that 100% of graduates are hired within one year from graduation, but also that they are assigned responsibility roles. This is the reward for their commitment and motivation.



---

# Contents

|          |                                                                          |           |
|----------|--------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Objectives and admission</b>                                          | <b>7</b>  |
| 1.1      | Objectives of the study program . . . . .                                | 7         |
| 1.2      | An inter-class program . . . . .                                         | 7         |
| 1.3      | Admission requirements . . . . .                                         | 8         |
| 1.4      | Pre-requisites . . . . .                                                 | 8         |
| <b>2</b> | <b>Program overview</b>                                                  | <b>11</b> |
| 2.1      | Study program . . . . .                                                  | 11        |
| 2.2      | Precedences . . . . .                                                    | 13        |
| 2.3      | Study plan . . . . .                                                     | 13        |
| <b>3</b> | <b>Teaching and service organization</b>                                 | <b>15</b> |
| 3.1      | Teachings . . . . .                                                      | 15        |
| 3.2      | International mobility: Erasmus+ and double degree . . . . .             | 16        |
| 3.3      | Service organization . . . . .                                           | 16        |
| 3.4      | Internship project and graduation . . . . .                              | 17        |
| <b>4</b> | <b>Relation to job market</b>                                            | <b>19</b> |
| <b>5</b> | <b>Relation to research activity</b>                                     | <b>21</b> |
| <b>A</b> | <b>Courses in English</b>                                                | <b>23</b> |
| A.1      | Compulsory subjects . . . . .                                            | 23        |
|          | <i>Business process modeling (295AA)</i> . . . . .                       | 23        |
|          | <i>Data mining (420AA)</i> . . . . .                                     | 24        |
|          | <i>Decision support systems (801AA)</i> . . . . .                        | 25        |
|          | <i>Optimization for data science (784AA)</i> . . . . .                   | 26        |
|          | <i>Statistics for data science (628PP)</i> . . . . .                     | 27        |
| A.2      | Elective subjects from the GR1 group . . . . .                           | 28        |
|          | <i>Advanced databases (641AA)</i> . . . . .                              | 28        |
|          | <i>Advanced laboratory of complex network analysis (0017A)</i> . . . . . | 29        |
|          | <i>Databases (765AA)</i> . . . . .                                       | 30        |
|          | <i>Distributed data analysis and mining (687AA)</i> . . . . .            | 31        |
|          | <i>Foundations of Computer Science (687AA)</i> . . . . .                 | 32        |
|          | <i>Geospatial analytics (783AA)</i> . . . . .                            | 33        |
|          | <i>Information retrieval (289AA)</i> . . . . .                           | 34        |
|          | <i>Machine Learning (654AA)</i> . . . . .                                | 35        |

|          |                                                                         |           |
|----------|-------------------------------------------------------------------------|-----------|
|          | <i>Programmatic advertising (634AA)</i> . . . . .                       | 36        |
|          | <i>Social network analysis (668AA)</i> . . . . .                        | 37        |
|          | <i>Text analytics (635AA)</i> . . . . .                                 | 38        |
|          | <i>Visual analytics (602AA)</i> . . . . .                               | 39        |
| A.3      | Elective subjects from the GR2 group . . . . .                          | 40        |
|          | <i>Advanced Statistical Learning (0106A)</i> . . . . .                  | 40        |
|          | <i>Financial Analysis and Performance Measurement (0003P)</i> . . . . . | 41        |
|          | <i>Fundamentals of business management (627PP)</i> . . . . .            | 42        |
|          | <i>Management practice (629PP)</i> . . . . .                            | 43        |
|          | <i>Strategic and competitive intelligence (787II)</i> . . . . .         | 44        |
| A.4      | Elective subjects from the GR3 group . . . . .                          | 45        |
|          | <i>Legal issues in data science (381NN)</i> . . . . .                   | 45        |
|          | <i>Model-Driven Decision-Making Methods (666AA)</i> . . . . .           | 46        |
|          | <i>Programming and Algorithms for Data Science (0104A)</i> . . . . .    | 47        |
| A.5      | Elective subjects from the Table 2.4 group . . . . .                    | 48        |
| <b>B</b> | <b>Corsi in Italiano</b>                                                | <b>49</b> |
|          | <i>Ingegneria del software* (271AA)</i> . . . . .                       | 50        |

---

# Objectives and admission

## 1.1 Objectives of the study program

The two-years graduate program in *Data Science and Business Informatics* has been designed to meet the constant demand for professionals with an interdisciplinary skills both in informatics and in business to satisfy the increasing demand by companies to compete using analytics and data science methods. The graduate program is focused on Business Intelligence and Data Science techniques to support decision making. The interdisciplinary competence covered by the Data Science and Business Informatics degree is intended to overcome the cultural divide between IT and management. In fact, as reported by several studies and publications, there is a shortage of trained professionals who can integrate the various skills and approaches necessary to overcome the traditional distrust of managers in involving computer professionals in decision-making. This is because computer professionals tend to be regarded as bearers of important but highly specialized knowledge, which may either seem difficult to apply or which has little relevance to the needs of organizations.

The professional profile of graduates in *Data Science and Business Informatics* is wider than that of traditional IT professionals, with specific expertise to support business activities and decision-making processes in operational information systems. The aim is to provide graduates with specific professional skills that lead to an effective dialogue with managers to support tactical and strategic activities. To achieve this goal, key expertise include *Business Intelligence* and *Data Science* (e.g., data warehousing, machine learning and artificial intelligence, data mining, business process modeling, big data analytics, visual analytics, text analytics). These technological skills are complemented by economic, statistical, and managerial skills in both operational, tactical and strategic activities rather. In fact, their economic and business skills enable graduates to understand the operational performance of an organization and the decision-making criteria adopted by the management, while the computer science skills enable them to translate this understanding into models using the latest IT technologies, in order to contribute significantly to an organization's decision-making and to the systematic innovation of products, processes, and services.

Finally, *Data Science and Business Informatics* graduates will have the necessary skills to access more advanced levels of university education, such as PhD programs in Artificial Intelligence, Data Science, Computer Science, etc.

## 1.2 An inter-class program

The Italian university system classifies study programs in [classes](#). Starting from A.Y. 2022/23, the graduate program in *Data Science and Business Informatics* is an inter-class program of LM-18 Informatics (the Computer Science class) e LM-DATA Data Science (the Data Science class). LM-18 Informatics is the classical Computer Science class, which allows, after graduation, to take the exam to access the Italian [professional order of Computer Engineers](#). Instead, LM-DATA Data Science is a fairly recent class of master programmes; while it does not allow for accessing the Italian professional order of Computer Engineers, it does allow for teaching of Mathematics or Physics (depending on the Bachelor study program) in Italian high-schools ().

At enrolment time, unless explicitly chosen by the student, the degree class is LM-18 (Computer Science). Students who intend to change the class to LM-DATA, within the first year, can do so by filling in the form

[available here](#) and sending it to the Students' Office (att. [Mrs Rosaria Mongini](#)).

There is only one curriculum, *independently from the classes students are enrolled in*, which includes an intensive study of in computer science topics, specifically data science subjects, as well as topics in mathematical optimization and statistics, and organization management. Students have a high freedom of choice in completing their curriculum. Thesis can be done as a research project or as an internship in one of the many companies that collaborate with the Department of Computer Science.

### 1.3 Admission requirements

Applicants must hold a first cycle degree in Computer Science (Italian class: L-31), Computer Engineering (L-8), Statistics (L-41), Economics (L-18 or L-33), Physics (L-30), Mathematics (L-35) or in the same classes according to Italian Law D.M. 509/1999. Students are admitted also if they hold a Bachelor degree with at least 40 ECTS credits in the following areas: Management and Economics (ECON-\*, IEGE-01/A), Informatics (INFO-01/A, IINF-05/A), Physics (PHYS/\*), Mathematics (MATH/\*), Statistics (STAT-\*). In the case of academic qualifications obtained abroad, exceptions may be made only with a resolution of the Admissions Committee, on the basis of the specific background of the candidate.

Applicants must also be fluent in English, e.g., holding a certificate equivalent to level **B2 or higher of CEFR**, or holding one of the following:

- at least 6 ECTS of ANGL-01/C “Lingua, traduzione e linguistica inglese”;
- proficiency certificate at level B2.1 rilasciato given by CLI (Centro Linguistico) of the University of Pisa;
- IELTS with score at least 5.5;
- TOEFL iBT with score at least 72;
- First Certificate in English;
- B2 certification given by any (public or private) authorized body;
- any other studies and experiences to be assessed by an interview.

Sufficient knowledge of English is required in practice as (almost) all lectures are in such language, hence students are advised not to underestimate this aspect; however, there is flexibility about the formal certification. Students with adequate knowledge but having not yet achieved a certification are still admitted, under the requirement that the certification will have to be eventually obtained before graduation (and possibly within the first year).

The Master degree does not have a maximum quota of admitted Italian or EU students, while there is a quota on the number of *extra-UE students* that can enroll. Usually, the application process for extra-UE students regarding intake of the Academic Year starting in September of year 20XX opens in December of the previous year (20[XX-1]) and closes around March of the year 20XX, in order for the students to have plenty of time to apply for their study visas (well-known to be a long and difficult process in many cases). Information on how to apply for the Master Programme can be found at:

<http://www.di.unipi.it/en/education/mds/enrollment-for-foreign-students>

Pre-applications will undergo a selection process, whereby requests will be ranked by means of the correspondence of the Bachelor studies to the intake requirements of the programme (see next section) and the quality of the study, measured, e.g., by means of the CGPA and on the quality of the graduating institution. For more information, follow the link above.

### 1.4 Pre-requisites

Basic knowledge on discrete mathematics, logics, computer programming, algorithmics, and data bases is required. Such topics are typically part of Bachelor programs in Computer Science or in Computer Engineering. Students with other Bachelor programs, if admitted, will learn such topics through one or more of the following elective subjects based on their specific Bachelor program:

- *Programming and Algorithms for Data Science* (12 ECTS);
- *Foundations of Computer Science* (6 ECTS);
- *Databases* (6 ECTS).



---

Furthermore, basic knowledge of the workings of industrial and business organizations will be required. This is given for students with Bachelors in Economics or Management, but typically not for students from the “hard sciences”: thus, the programme offers the course of *Fundamentals of business management* (9 ECTS) that is required for all students without the corresponding prerequisites.



---

# Program overview

The Master programme is offered by the Department of Computer Science, and it has the following structure:

- 48 ECTS credits from the *Informatics* area;
- 15 ECTS credits from the *Mathematics and Statistics* area;
- 9 ECTS credits from the *Business Economics and Business Law* areas;
- 12 ECTS credits from the *Business Economics, Business Law, Mathematics and Informatics* areas;
- 9 ECTS credits of elective subjects;
- 27 ECTS credits for a final thesis, which can be associated with an internship in a public institution or in a private company, both in Italy and abroad.

The effort for each subject is given in ECTS, each of which consists on average of 25 hours of study, typically as 8 hours of attending lectures plus 17 hours of individual study. Courses are organized in two semesters per year. Each course is taught in a specific semester only, except for one annual course (Programming and Algorithms for Data Science) taught over two semesters.

## 2.1 Study program

The study program depends on the student's Bachelor degree, but in all cases it will be designed to give an interdisciplinary expertise in informatics, data science and business. Most of the courses are taught in English; however, a few ones are taken by other programmes and are taught in Italian. These can be recognised, in the list below, by the Italian title and by being marked with an asterisk (\*).

### Courses from the *Informatics* area (48 ECTS)

- **Data mining** (INFO-01/A ECTS 12 DM 420AA)
  - Module I: Data mining: fundamentals (6 ECTS 1 Sem.)
  - Module II: Data mining: advanced topics and applications (6 ECTS 2 Sem.)
- **Business process modeling** (INFO-01/A ECTS 6 BPM 295AA 1 Sem.)
- **Decision support systems** (INFO-01/A ECTS 12 DSS 801AA)
  - Module I: Decision support databases (6 ECTS 1 Sem.)
  - Module II: Laboratory of data science (6 ECTS 1 Sem.)
- **Elective courses from GR1 group from Table 2.1 (18 ECTS)**

| Course                                          | Description |      |       |       |      |
|-------------------------------------------------|-------------|------|-------|-------|------|
|                                                 | Area        | ECTS | Abbr. | Code  | Sem. |
| Advanced databases                              | INFO-01/A   | 9    | ADB   | 641AA | 2    |
| Advanced laboratory of complex network analysis | INFO-01/A   | 6    | ALC   | 0017A | 1    |
| Databases                                       | INFO-01/A   | 6    | DB    | 765AA | 2    |
| Distributed data analysis and mining            | INFO-01/A   | 6    | DDAM  | 687AA | 1    |
| Foundations of Computer Science                 | INFO-01/A   | 6    | FCS   | 0105A | 1    |
| Geospatial analytics                            | INFO-01/A   | 6    | GSA   | 783AA | 1    |
| Information retrieval                           | INFO-01/A   | 6    | IR    | 289AA | 1    |
| Machine learning                                | INFO-01/A   | 9    | ML    | 654AA | 1    |
| Programmatic advertising                        | INFO-01/A   | 6    | PRV   | 634AA | 1    |
| Social network analysis                         | INFO-01/A   | 6    | SNA   | 668AA | 2    |
| Text Analytics                                  | INFO-01/A   | 6    | TX    | 635AA | 1    |
| Visual analytics                                | INFO-01/A   | 6    | VA    | 602AA | 2    |

**Table 2.1** GR1 Group of choices: the *Informatics* area

**Courses from the *Mathematics and Statistics* area (15 ECTS)**

- **Optimization for data science** (MATH-06 CFU 6 ODS 784AA 1 Sem.)
- **Statistics for data science** (STAT-01/A CFU 9 SDS 628PP 2 Sem.)

**Courses from the *Business Economics and Business Law* areas (9 ECTS)**

- **Elective courses from GR2 group from Table 2.2 (9 ECTS)**

| Course                                         | Description          |      |       |       |      |
|------------------------------------------------|----------------------|------|-------|-------|------|
|                                                | Area                 | ECTS | Abbr. | Code  | Sem. |
| Advanced Statistical Learning                  | STAT-01/A, INFO-01/A | 6    | ASL   | 0106A | 2    |
| Change Management and Digital Transition       | ECON-08/A            | 9    | CMDT  | 0033P | 1    |
| Computer Science Law                           | GIUR-01/A            | 6    | CSL   | 058NN | 1    |
| Data Analytics for Planning and Control        | ECON-06/A            | 9    | DAPC  | 0040P | 1    |
| Financial analysis and performance measurement | ECON-06/A            | 9    | FAP   | 0003P | 1    |
| Fundamentals of business management            | ECON-06/A            | 9    | FBM   | 627PP | 1    |
| Management practice                            | ECON-07/A            | 6    | MP    | 629PP | 2    |
| Strategic and competitive intelligence         | IEGE-01/A            | 6    | SCI   | 787II | 2    |

**Table 2.2** GR2 Group of choices: the *Business Economics and Business Law* areas

**Courses from the *Business Economics, Business Law, Mathematics and Informatics* areas (12 ECTS)**

- **Elective courses from GR2 group from Table 2.2 and/or from GR3 group from Table 2.3 (12 ECTS)**
- Note that the 9 ECTS courses from GR2 group selected in the previous section do not contribute to the 12 ECTS of those in this section, that must therefore be additional to those.

| Course                                      | Description |      |       |       |      |
|---------------------------------------------|-------------|------|-------|-------|------|
|                                             | Area        | ECTS | Abbr. | Code  | Sem. |
| Algorithmic Game Theory                     | MATH-06/A   | 6    | AGT   | 759AA | 2    |
| Ingegneria del software*                    | INFO-01/A   | 6    | IS    | 271AA | 1    |
| Legal issues in data science                | GIUR-11/A   | 6    | LDS   | 381NN | 2    |
| Model-Driven Decision Making Methods        | MATH-06/A   | 6    | MDD   | 666AA | 1    |
| Programming and Algorithms for Data Science | INFO-01/A   | 12   | PADS  | 0104A | 1/2  |

**Table 2.3** GR3 Group of choices: the *Informatics, Mathematics, and Business Law* areas

---

## Courses from elective subjects (9 ECTS)

- One or two courses among the ones from groups GR1, GR2, GR3 and GR4 (see Table 2.4) will have to be selected to reach 9 ECTS. Suggested courses depend on the Bachelor degree program of the student. Courses from other Master programmes can be considered as well.

| Course                                      | Description |      |       |      |
|---------------------------------------------|-------------|------|-------|------|
|                                             | Area        | ECTS | Code  | Sem. |
| Artificial Intelligence fundamentals        | INFO-01/A   | 6    | 643AA | 1    |
| Continual learning                          | INFO-01/A   | 6    | 791AA | 2    |
| ICT risk assessment                         | INFO-01/A   | 9    | 303AA | 1    |
| Peer to peer systems and blockchains        | INFO-01/A   | 9    | 261AA | 2    |
| Intelligent systems for pattern recognition | INFO-01/A   | 9    | 760AA | 2    |

**Table 2.4** GR4 Group of choices: suggested elective subjects

## 2.2 Precedences

There is no formal precedence between courses. However, the following order should be respected to be able to attend subjects with profit:

- for **Advanced lab of complex network analysis** to have attended: **Social network analysis**;
- for **Decision support systems** to have attended: **Data mining**.
- for **Distributed data analysis and mining** to have attended: **Data mining**;
- for **Geospatial analytics** to have attended: **Data mining**;
- for **Information retrieval** to have attended: **Programming and Algorithms for Data Science** (if in the study plan);
- for **Model-driven decision making methods** to have attended: **Optimization for data science**;
- for **Advanced Statistical Learning** to have attended: **Statistics for data science**;
- for **Statistics for data science** to have attended: **Optimization for data science**;
- for **Strategic and competitive intelligence** to have attended: **Fundamentals of business management** (if in the study plan).

These guidelines are especially relevant for students enrolled late in the first semester.

## 2.3 Study plan

Students are required to submit the study plan within one month from enrollment. Forms for submitting the study plan and draft study plans specifically tailored to several different Bachelor degrees are available at:

<https://didattica.di.unipi.it/en/master-programme-in-data-science-and-business-informatics/study-plans>

The study plan can be updated at any point in time, using the appropriate form available at the link above.



---

# Teaching and service organization

The website of Master degree programme contains the updated information on the overall organization as well as news and announcements:

<https://www.di.unipi.it/en/education/mds>

Some information related to course and services offered by the Department of Economics and Management are described at the website:

<https://www.ec.unipi.it/didattica>

## 3.1 Teachings

### Academic calendar, timetable, and rooms

The academic calendar is organized over two semesters (September-December and February-May), including each at least 12 weeks of teachings. The courses from Table 3.1 may have a slightly different organization from the others, because they are offered by the study programmes of the Department of Economics and Management.

| Course                                   |
|------------------------------------------|
| Change Management and Digital Transition |
| Computer Science Law                     |
| Data Analytics for Planning and Control  |

**Table 3.1** Courses offered by the Department of Economics and Management.

Timetable of courses is published on the website before the beginning of the semester. Teaching rooms are located as follows:

- for the courses from Table 3.1 at the Department of Economics and Management, via C. Ridolfi 10, Pisa;
- for all other courses, at Polo Didattico L. Fibonacci, Largo B. Pontecorvo 3, Edificio B, Pisa.

Notice that the time slots for the courses from Table 3.1 (8:45-10:15, 10:30-12:00, 12:15-14:00, 14:15-15:45, 16:00-17:30, 17:45-19:15) are different from those of all the other courses (9-11, 11-13, 14-16, 16-18).

### Attendance

No course requires mandatory attendance.

### Course program and teaching material

The VALUTAMI web portal provides a list and several information on courses:

[esami.unipi.it](https://esami.unipi.it)

For each course, there is a description of the learning objectives and course program. Moreover, there is a link to the course web page, maintained by the teacher of the course, with additional information, such as: calendar of lessons, slides, teaching material, exercises, audio-video recordings, etc.

The official final program of a course is available at the website <http://unimap.unipi.it>, by searching for the surname of the teacher, then the link “Attività didattica” and then the link “Registro delle lezioni”.

### Exams and mid-terms

Exams consist typically of a written test and/or a project and an oral test. Sometimes, the written test can be passed during the semester through mid-term tests, typically one at the mid of the semester and one at the end of the semester. The academic calendar states the periods when mid-terms can take place. The grades for passing an exam are on a scale from 18 to 30 cum laude. The given grade can be declined by the student: in such a case, the exam must be repeated from scratch. The enrollment to an exam session (“appello”) must be done on the VALUTAMI web portal <https://esami.unipi.it>. For a few categories of students, there are two additional exam sessions (“appelli straordinari”), scheduled around the mid of each semester. Such categories are: students enrolled from third year on (“fuori corso”), parents with children below eight year old, pregnant students, working students. See the following links

<https://didattica.di.unipi.it/disciplina-studenti-lavoratori>  
<https://didattica.di.unipi.it/appelli-straordinari>

for details about the admission criteria and procedure to apply for the additional exam sessions.

### Student questionnaire

In the last weeks of each semester, students must fill a questionnaire for each course attended during the semester (for courses over the whole academic year, this is required only in the second semester). Questionnaire can be filled on the web portal VALUTAMI <https://esami.unipi.it>. When enrolling to an exam session, the system will check if the questionnaire has been filled. If not, before enrolling, the student will be forced to fill the questionnaire. The answers are anonymized and no link is left between the student and the answers. The questionnaire is extremely relevant to understand the feelings of the students about the course contents, materials, and teacher(s). They will be thoroughly considered for taking actions to improve the course in the following years.

## 3.2 International mobility: Erasmus+ and double degree

Students can apply for the Erasmus+ programme for a period (one semester, one year) in a European university both for taking exams or for doing the thesis. Incoming Erasmus+ students also regularly come to Pisa to attend the courses of the Master degree programme in Data Science and Business Informatics.

A double degree program consists of a study program designed by two Universities. The graduated student will have two titles of study, given by each of the Universities. Students can enroll to a double degree program with the University of Paris-Dauphine (France), offering the *Master in Informatique des Organisations (parcours Informatique pour la Décision de la 1ère année de Master et parcours MIAGE Informatique Décisionnelle de la 2ème année de Master)*. There is a quota of five students per year admitted to the double degree program.

For more information:

<https://didattica.di.unipi.it/en/master-programme-in-data-science-and-business-informatics/studying-abroad-msc-in-data-science>

## 3.3 Service organization

### Office hours and tutoring

Contacts, office addresses and office hours (weekly or on appointment) of teachers are available at the website <https://unimap.unipi.it/cercapersone/cercapersone.php> or directly at the web pages of courses. Students are strongly suggested to take advantage of office hours for clarifying any doubt about the topics of a course and the contents of each class, or for any other inquiry about a course.

By the first semester of each year, each newly enrolled student is assigned a tutor teacher, who can be contacted for general inquiries about teaching. See



---

<https://didattica.di.unipi.it/laurea-magistrale-in-data-science-and-business-informatics/servizio-di-tutorato-3>

for details about the assignment process and the duties of tutors and students. For privacy reasons the list of assigned tutors cannot be published on the web site; each student can ask her/his tutor by e-mail to <mailto:rosaria.mongini@unipi.it>.

In the first semester, there also will be an au-pair tutoring service, in which students of the second year provide information and guidance to newly enrolled students.

### **Computer labs, Wi-Fi and software licences**

The computer labs at the Polo Didattico L. Fibonacci can be accessed outside of the scheduled classes. For logging in to computers or for connecting to the Wi-Fi (name “Unipi”), students can use the credentials of Alice. For other IT services offered by the University of Pisa, consult <https://it.unipi.it>.

### **Student secretariats and Alice portal**

The student secretariat is located at Largo B. Pontecorvo 3, Edificio E. It is the primary reference for enrolling (through the International Office), for taxes, for obtaining formal certificates, for applying for graduation. Most of such services can be done online on the student web portal Alice: <https://www.studenti.unipi.it>.

*Important:* Alice is not synchronized with the study plan approved (which is the sole official). Some passed exams may appear as “extra” (“sovrannumerario” in Italian). If you need to remove the “extra” label (e.g., because you are requested so by the DSU, the office for scholarships), please send an email to

<mailto:scienze.segreteriastudenti@unipi.it>, with <mailto:datascience@di.unipi.it> in cc

and attaching the approved study plan.

The didactic unit of the Master Programme is located at Dipartimento di Informatica, Largo B. Pontecorvo 3, Edificio C, II piano. It is the primary reference for the admin papers for doing an internship, and for applying for graduation (you have to apply both online and at the didactic unit).

### **Centro Linguistico Interdipartimentale**

The *Centro Linguistico Interdipartimentale* (CLI, <https://www.cli.unipi.it>) offers language courses, including course of Italian for foreigner students and the B2 certification for students who does not already have it at enrolling time.

### **Students with disabilities**

Students with disabilities can be supported in several ways. Please contact *Ufficio Servizi per l'Integrazione di studenti con Disabilità* (USID, <https://www.unipi.it/index.php/usid>). The USID offices offer technical and IT aids, tutoring, and other support facilities.

Students with Specific Learning Disorders can contact the [sportello DSA](#), which offers assistance, tutoring, and support to them.

The [Servizio di Ascolto e Consulenza per Studenti Universitari](#) aims at supporting any student who need psychological assistance while studying at the University of Pisa.

### **Questionnaire for the evaluation of services and organization**

Students shall fill once a year a questionnaire about the quality of services and organization. The questionnaire is available at the VALUTAMI portal <https://esami.unipi.it>. The VALUTAMI portal will prompt for filling the questionnaire (if not already filled) when registering to an exam session.

## **3.4 Internship project and graduation**

### **Thesis and Internship project (“Progetto formativo”)**

Graduation is subject to the completion of a Master Thesis, weighted at 27 ECTS; basically, it's the whole fourth semester of the course, i.e., roughly 6 months of work. The Thesis has to result in a written document of roughly 80-100 pages (with large discretion) properly documenting the work done; the Theses are deposited at <https://etd.adm.unipi.it>; search for “Corso di Studi” = “data science” to get the list of past ones. The Thesis can be done in two ways:

- “Academic” Thesis: each student, when having only roughly one or two exams left at most still to give, can autonomously contact any lecturer of the course, or other professors/researchers of the Department, the University of Pisa at large, and related research institutions like [CNR Pisa](#) (in particular, the ISTI and IIT departments), [Scuola Sant’Anna](#) and [Scuola Normale Superiore](#) to ask for available research projects in subjects related with the Master Course (at large). Upon finding a match, the Thesis can start right away with no formalities if the professor/researcher is a lecturer of the course, otherwise a mail to <mailto:antonio.frangioni@unipi.it> is advised. The Thesis will terminate when the agreed work has been carried out.
- An internship project (“Progetto formativo” in Italian), consisting of an agreement among the Department of Computer Science of the University of Pisa, a student preparing for doing the thesis, and a company or public institution, in Italy or abroad, that is willing to host the student for some period and with shared objectives. Students become part of a project team in a company project focusing on topics of Data Science and Business Informatics that have the potential to provide enough material for the Master thesis document. The student is supervised by a company tutor and by an academic tutor, who will act as the thesis supervisor as well. See <https://didattica.di.unipi.it/en/master-programme-in-data-science-and-business-informatics/internship> for the rules of the internship, a list of companies with an agreement, and the list of available proposals. It is common and welcome that students autonomously find out new companies/institutions not already in the list: stipulating the appropriate agreement is quick and easy. The internship can only start when the student only has one exam left at most still to give, and will have to last around 670-700 hours—comprising the time to write the document—as testified by a daily timesheet co-signed by the student and the company tutor.

Internship projects are a good way to make practical work experience and are often useful to help job finding.

### **Graduation (“Laurea”)**

Deadlines for graduation and instructions for applying are available at the web site of the Master programme. See

<https://didattica.di.unipi.it/en/master-programme-in-data-science-and-business-informatics/graduation-2>

---

## Relation to job market

The current phase of digital transformation has triggered a process of innovation in organizations that builds on investing in new technologies, in their integration with the business processes, and in the constant increase of data available for analyses. This transformation is profoundly changing the information and communication technology (ICT) skills required by public institutions and private companies. Such skills will be more and more oriented towards multidisciplinary, analytics, and process and sectorial skills ([Assintel Report 2021](#) “ICT and Digital Market in Italy”).

The recent reports “[The Future of Jobs](#)” published by the World Economic Forum, based on a global survey in the many production sectors, systematically identify “big data” and related technologies as one of the disruptive factors of change, and in the “data scientist/data analyst” one of professional figures that emerge as critical and indispensable in many sectors. The demand for professionals capable of reading and managing data, of solving advanced problems with an analytical approach, of exploiting innovation itself to improve the quality of the business, is estimated to grow more and more in the following years.

The authoritative American research company Gartner, in its periodic reports on [Analytics and Business Intelligence platforms](#), claims that the market for Business Intelligence and Analytics platforms will remain one of the fastest growing software segments. Its growth is guaranteed by the process of digital transformation underway, from the increase in investments in analytics by companies and by the increase in the availability of data. The Gartner reports confirm the trend, already identified in the past, (see: T.H. Davenport and G.C. Harris, *Competing on Analytics: The New Science of Winning*, Harvard Business School Press, Boston 2007, and T.H. Davenport, G.C. Harris and R. Morison, *Analytics at Work*, Harvard Business School Press, Boston 2010) on the ability and willingness of modern companies to collect big data and to invest in strategies based on the analysis of such data with the Business Intelligence tools. In a famous article on Harvard Business Review, T.H. Davenport defined the data scientist “[the Sexiest Job of the 21st Century](#)”.

The latest analyses of the Italian context by the [Big Data and Business Analytics Observatory](#) of the School of Management of the Polytechnic of Milan show that, despite the slowdown due to the pandemic, 96% of large companies continue to perform activities to improve the collection and valorisation of data and 42% have taken action, in terms of projects and skills, in the Advanced Analytics field. Organizations understand well that Business Intelligence tools and methods allow to create applications that help managers better understand their business and discover new possibilities for intervention to compete in today’s world. As pointed out by the Observatory itself in a 2007 study (*Business Intelligence: a look at the future*), and in the literature (see e.g., the book by T. Burelli, A. Marzona, M. Pighin, *From intuition to knowledge*, Aracne, Rome, 2007), to make the most of the potential of decision support systems, companies must invest not only in technology, but also on graduates with specialized skills in computer science, economics, business management, data mining and in optimization methods. The *Data Science and Business Informatics* degree program bases its educational offering on the objective of transmitting to its own students exactly these skills, in line with expectations of the labour market.

The *Data Science and Business Informatics* programme promotes and encourages master’s theses developed in collaboration with private companies and public institutions, thus favoring the rapid employability of graduates. From the results of occupational surveys of graduates conducted as part of the initiative STELLA (until 2014) and by [AlmaLaurea](#) (since 2015), it turns out that the employment rate of graduates in *Data Science and Business Informatics* is among the highest of the University of Pisa master programmes.



---

## Relation to research activity

The Master degree programme is run by the *Department of Computer Science* of the University of Pisa. The Department of Computer Science employs 23 full professors, 32 associate professors, 23 researchers. They cover the scientific disciplinary sectors of Computer Science, Operational Research, Numerical Analysis, and Business Management. The Department of Computer Science also runs a PhD in Computer Science, offered in collaboration with the University of Florence and the University of Siena, and the National PhD programme in Artificial Intelligence and Society, jointly between the University of Pisa and CNR, Scuola Superiore Sant'Anna, Scuola Normale Superiore, IMT Lucca School, and several other universities.

The *Data Science and Business Informatics* programme also benefits from the contribution of teachers:

- of the *Department of Economics and Management*, in particular for the sharing of some courses in the Business Management sector;
- of the *National Research Council*, in particular of the *Institute of Science and Technology of Information* of Pisa;
- of the *Department of Energy Engineering, Systems, Territory and Construction*, for courses of the sector of Managerial Engineering;
- of the *Department of Law*, for courses of the sector of Private Law;
- of the *Scuola Superiore S. Anna*, for courses of the sectors of Comparative Private Law and of Economics and Business Management.

The teachers of the Master programme are engaged in research activities that are coherent, relevant and useful with respect to the objectives of the courses they taught. In particular, the following research areas are covered:

- Operational information systems: theories, techniques, languages, architectures and systems for databases and XML.
- Management information systems: theories, techniques, data science languages, data architectures and systems, business intelligence, data warehousing, data mining, big data, text mining.
- Machine learning and artificial intelligence: theories, algorithms and applications of machine learning (deep learning, reinforcement learning, learning for structured/complex data, patterns recognition, natural language processing), ethical issues in artificial intelligence.
- Business processes and services on the web: theories, techniques, languages, architectures and systems for business processes, planning and coordination of inter-company services.
- Models and methods for quantitative-based decisions support systems, employing mathematical models and the corresponding solution algorithms, for applications such as logistics systems (production, transportation, distribution, . . . ), energy systems, finance, health, and many others.
- Economics and management: economic theory of demand, economic equilibrium, international economics, planning and control, competitive intelligence.

The multidisciplinary nature of the teachers with a mix of cutting-edge cultural, scientific and technological skills allows for guiding students along training paths to highly qualified professionals that are highly sought after in labour market.



---

# Courses in English

This appendix reports the syllabus of the courses offered in English. Details on each course are available at the Course Catalogue:

<https://unipi.coursecatalogue.cineca.it/corsi/2026/11357;anno=2026;corso=11357>

and the web page of the course.

## A.1 Compulsory subjects

### **Business process modeling (295AA) (6 ECTS)**

*Semester:* 1

*Contact Person:* Prof. Roberto BRUNI (<mailto:roberto.bruni@unipi.it>)

#### *Objectives*

The course presents techniques for Business Analytics according to the process-driven view of Business Process Modeling. It presents the main concepts and problematic issues related to the process management, where processes are understood as workflow over some basic activities, and to show some of the languages, conceptual models and tools that can help to handle the main problems in a proper way. During the course, the students will become acquainted with the technical terminology of the area, with several rigorous models that can be used to structure and compose processes, with the logical properties that such processes can be required to satisfy and with specific analysis and verification techniques. Moreover they will be given the possibility to experiment with some advanced tools for the design and analysis of business processes.

#### *Syllabus*

- Introduction to Key Issues in Business Process Management.
- Terminology and Classification.
- Process Modeling. Conceptual Models and Levels of Abstraction.
- Rigorous Workflow Models: Petri Nets and Workflow Nets.
- Tool-supported Workflow Design and Analysis: Experimentation with Integrated Tools for Business Process Design, Analysis and Verification.

## **Data mining (420AA) (12 ECTS)**

*Semester:* 1, 2

*Contact Person (1st sem.):* Prof. Dino PEDRESCHI (<mailto:dino.pedreschi@unipi.it>)

*Contact Person (2nd sem.):* Prof. Riccardo GUIDOTTI (<mailto:riccardo.guidotti@unipi.it>)

### *Objectives*

Recent tremendous technical advances in processing power, storage capacity, and interconnectivity are creating unprecedented quantities of digital data. Data mining, the science of extracting useful knowledge from such huge data repositories, has emerged as an interdisciplinary field in computer science. Data mining techniques have been widely applied to problems in industry, science, engineering and government, and it is believed that data mining will have profound impact on our society. The course is divided into two modules. The first presents an introduction to the basic concepts of data mining and the knowledge discovery process, and associated analytical models and algorithms. The second module provides an account of advanced techniques for analysis and mining of novel forms of data, and the main application areas and prototypical case studies.

### *Syllabus*

#### Module 1: Foundations

- Concepts of Data Mining and the Knowledge Discovery Process.
- Data Preprocessing and Exploratory Data Analysis.
- Frequent Patterns and Associations Rules.
- Classification: Decision Trees and Bayesian Methods.
- Cluster Analysis: Partition-based, Hierarchical and Density-based Clustering.
- Experiments with Data Mining Toolkits.

#### Module 2: Advanced topics and applications

- Mining Time-Series and Spatio-Temporal Data.
- Mining Sequential Data, Mining Large Graphs and Networks.
- Advanced Association, Correlation and Frequent Pattern Analysis.
- Advanced Classification, Cluster Analysis and Outlier Detection.
- Data Mining Languages, Standards and System Architectures.
- Ethical aspects of data mining.
- Privacy-Preserving Data Mining.
- Applications: Retail Industry, Marketing, CRM, Telecommunication Industry, Financial Data Analysis, Risk Analysis, Fraud Detection, Mobility and Transportation, Public Administration and Health.



---

## Decision support systems (801AA) (12 ECTS)

*Semester:* 1

*Contact Person (Module I):* Prof. Salvatore RUGGIERI (<mailto:salvatore.ruggieri@unipi.it>)

*Contact Person (Module II):* Prof. Anna MONREALE (<mailto:anna.monreale@unipi.it>)

### *Objectives*

The course presents the main methodological and technological approaches to the design and implementation of decision support systems based on business intelligence (data warehousing, data mining, data science). The first module covers themes such as conceptual and logical Data Warehouses design, data analysis using analytic SQL, algorithms for selecting materialized views, data warehouse systems technology (indexes, star query optimization, physical design, query rewrite methods to use materialized views). The second module presents technologies and systems for data access, for building and analyzing data warehouses, for reporting, and for knowledge discovery in databases. The accent of the module is on the use of tools and on the analysis of application problems by means of non-trivial samples and case studies.

### *Syllabus*

#### Module I: Decision support databases

- Information systems and computer-based information systems in organizations.
- Decision Support System Based on Data Warehouses.
- Data Models for Data Warehouses and On-line Analytical Processing.
- Conceptual and logical design in Data Warehouses.
- Algorithms for Selecting Materialized Views.
- Data Warehouse Systems Technology: Indexes, Star Query Optimization,
- Physical Design, Query Rewrite Methods to Use Materialized Views.
- Case studies.

#### Module II: Laboratory of Data Science

- Introduction: Tools for data science and Business Intelligence.
- Data Access. Location, Format and API for Accessing Data in Text Files. Standards for Data Connectivity.
- Extract Transform and Load. Tools for ETL. Case studies.
- Data Warehousing and OLAP. Tools for Dimensional Modeling. Case Studies.
- Tools for Reporting and Multidimensional Browsing. Case Studies.
- Data Mining. Tools for Knowledge Discovery. Case Studies.

## **Optimization for data science (784AA) (6 ECTS)**

*Semester:* 1

*Contact Person:* Prof. Antonio FRANGIONI (<mailto:antonio.frangioni@unipi.it>)

### *Objectives*

The course aims at familiarising the students with the mathematical optimization methodologies underpinning many Data Science approaches, as well as with their use in combination with Data Science techniques to address relevant practical problems. The course will therefore carefully balance the methodological aspect, i.e., the theory of constrained and unconstrained optimization and of the corresponding solution algorithms, and the applicative aspect, i.e., the use of these methodologies to address Data Science issues.

### *Syllabus*

- Linear algebra and calculus background (on a need-to-know basis).
- “Simple” unconstrained optimization, univariate and multivariate.
- Univariate continuous optimization.
- Multivariate continuous unconstrained optimization (smooth or not).
- Constrained continuous optimization theory.
- Convexity and Duality (Lagrangian, linear, quadratic, conic, ...).
- Some numerical methods for continuous constrained optimization.
- Sparse hints to Mixed-Integer optimization models and software tools.
- Sparse hints to Data Science applications.

---

## Statistics for data science (628PP) (9 ECTS)

*Semester:* 2

*Contact Person:* Prof. Salvatore RUGGIERI (<mailto:salvatore.ruggieri@unipi.it>)

### *Objectives*

The course introduces the fundamental concepts and techniques of probability and inferential statistics essential for data analysis and data science. After consolidating knowledge in descriptive statistics and probability theory, it focuses on key methods in estimation theory and hypothesis testing. The objective is to provide a solid foundation for validating core data analysis tasks and machine learning models. Theoretical concepts are complemented by practical exercises and project work using the R programming language.

### *Syllabus*

- Self-contained review on probability theory, random variables, distances between distributions and convergence theorems for sequences of random variables.
- Exploratory data analysis: graphical and numerical summaries.
- Basic statistical models.
- The bootstrap method.
- Estimation: unbiased estimators, efficiency and mean squared error, maximum likelihood, expectation maximization.
- Least squares estimation and regression.
- Confidence intervals and hypotheses testing.
- Distribution fitting.
- Resampling methods.
- Categorical data and inference for contingency tables.
- Classifier error rate estimation.

## A.2 Elective subjects from the GR1 group

### Advanced databases (641AA) (9 ECTS)

*Semester:* 2

*Contact Person:* Prof. Giorgio GHELLI (<mailto:giorgio.ghelli@unipi.it>)

*Web Page:* <https://elearning.di.unipi.it/enrol/index.php?id=986>

#### *Objectives*

The course provides advanced technical knowledge of the main issues related to the implementation and performance optimization of both classical centralized relational database systems for operational and OLAP processing and of recent advances in non-relational data models (columnar, document, key-value, graph) and scalable distributed architectures.

#### *Syllabus*

- Internals of relational database management systems.
- Optimizations of Data Warehousing management systems and On-Line Analytical Processing.
- Extract-Transform-Load and query/reporting in OLAP systems.
- Beyond SQL: NoSQL data management systems for big data.
- Distributed data processing and the Map-Reduce paradigm.

---

## **Advanced laboratory of complex network analysis (0017A) (6 ECTS)**

*Semester:* 1

*Contact Person:* Prof. Giulio ROSSETTI (<mailto:giulio.rossetti@isti.cnr.it>)

*Contact Person:* Prof. Barbara GUIDI (<mailto:barbara.guidi@unipi.it>)

### *Objectives*

Delving deep into the intricacies of complex systems—be they social, biological, or technological—is vital for accurately modeling and effectively tackling the pressing issues that define our world today. Take, for instance, the need to diminish polarization and radicalization in online discussions, predict high-frequency financial transactions, or understand the dynamic interactions among proteins. Building upon the theoretical groundwork laid in “Social Network Analysis,” this laboratory course aims to equip students with the tools to construct robust pipelines for analyzing complex systems derived from real-world data and represented as graphs. It will introduce key methodologies for data collection and preprocessing for graph analysis, delve into models designed to enhance the graph’s descriptive capabilities and practical applications, and explore techniques for evaluating experimental outcomes. The course emphasizes hands-on learning, empowering students to gain practical insights into these concepts through the utilization of advanced Python libraries.

### *Syllabus*

- Formulating Hypotheses.
- From Simple Graphs to Advanced Models.
- Network Sampling .
- Data Collection: API & Web Scraping.
- Graph Transformation: Backboning.
- Enriching Topologies: Feature-rich modeling.
- How to Validate: statistical significance of network-based analyses.

**Databases (765AA) (6 ECTS)**

*Semester:* 2

*Contact Person:* Prof. Giorgio GHELLI (<mailto:giorgio.ghelli@unipi.it>)

*Objectives*

The management of information is the main use of computers in organizations of all types and sizes. Information management is mostly based on data base technology. The aim of the course is to present the features of these systems, in particular the relational ones, their architecture and the principles they are inspired by, from the point of view of application designers.

*Syllabus*

- Data base and database management system: definition and functionalities.
- Database design: conceptual modeling using the object-oriented data model.
- The relational data model.
- Mapping of conceptual schemas onto relational logical schema.
- The SQL language, with a special emphasis on the query sublanguage and its relationship with first order logic.
- Theory of relational database normalization.
- Database implementation: access plans and transaction management.
- NoSQL systems.

---

## **Distributed data analysis and mining (687AA) (6 ECTS)**

*Semester:* 1

*Contact Person:* Prof. Roberto TRASARTI (<mailto:roberto.trasarti@isti.cnr.it>)

### *Objectives*

Mining with big data or big data mining has become an active research area. Running current analytical methodologies and software tools on a single personal computer cannot efficiently deal with very large datasets. Distributed computing platforms are a scalable solution for big data mining, obtained by dividing a large problem into smaller ones that are concurrently solved by many single processor/machine. This course aims at teaching the basic theoretical concepts behind the MapReduce distributed computing paradigm, and Hadoop in particular, and at building expertise in the practical usage of high performance computing tools for data engineering, analysis and mining. In particular the students will learn how the classical data mining algorithms can be applied on Big Data using Hadoop (Spark). Real (and open source) datasets will be used to present examples and to let the students build their own projects. Half of the lessons will consists of practice (Lab), and half of lectures.

### *Syllabus*

- Motivations: Distributed Data Mining in a Big Data Scenario.
- Recall parallel and distributed computing notions.
- Introduction to Hadoop.
- Hadoop Ecosystem.
- Interacting with HDFS (LAB).
- Map-Reduce Programming Patterns.
- Recall Python programming (LAB).
- Basic Spark (LAB).
- Data Analysis with Spark (LAB).
- Data Mining and Machine Learning with Spark (LAB).
- SparkSQL (LAB).
- Example on how to prepare a project.
- Real Case Studies.

## Foundations of Computer Science (0105A) (6 ECTS)

*Semester:* 1

*Contact Person:* Prof. Laura SEMINI (<mailto:laura.semini@unipi.it>)

### *Objectives*

The course provides the mathematical and logical foundations of computer science, with a balanced approach between theoretical rigor and applications to computer science and data science. The aim is to develop formal reasoning, mathematical modeling skills, and the ability to analyze computational problems. No mandatory prerequisites: basic mathematical knowledge is sufficient.

### *Syllabus*

- Mathematical logic: propositional logic and proof techniques; predicate logic and the use of quantifiers.
- Discrete structures: sets and basic set operations; multisets and fuzzy sets; sequences and functions; set cardinality.
- Number theory and representations: congruences modulo  $n$ ; prime numbers and greatest common divisors; the Euclidean algorithm; integer representations and base conversion; introduction to undecidable problems and hash functions.
- Induction and recursion: simple and strong mathematical induction; well-ordering principle; recursive definitions; structural induction; recursive functions and algorithms.
- Combinatorial counting: fundamental counting principles; permutations and combinations with and without repetition; the pigeonhole principle; the principle of inclusion–exclusion.
- Relations and orderings: binary relations and their properties; representations using ordered pairs, matrices, and digraphs; partial and total orders; Hasse diagrams; minimal and maximal elements; compatible orders and topological ordering.
- Equivalence relations: equivalence relations, equivalence classes, and partitions; applications.
- Linear algebra: vectors, matrices, and their main applications in computer science.



---

## Geospatial analytics (783AA) (6 ECTS)

*Semester:* 1

*Contact Person:* Prof. Luca PAPPALARDO (<mailto:luca.pappalardo@isti.cnr.it>)

### *Objectives*

The analysis of geographic information, such as those describing human movements, is crucial due to its impact on several aspects of our society, such as disease spreading (e.g., the COVID-19 pandemic), urban planning, well-being, pollution, and more. This course will teach the fundamental concepts and techniques underlying the analysis of geographic and mobility data, presenting data sources (e.g., mobile phone records, GPS traces, geotagged social media posts), data preprocessing techniques, statistical patterns, predicting and generative algorithms, and real-world applications (e.g., diffusion of epidemics, traffic simulation, urban dynamics). The course will also provide a practical perspective through the use of advanced geanalytics Python libraries.

### *Syllabus*

- Spatial Reference Systems.
- Data formats.
- Trajectory and Flows.
- Spatial Tessellations.
- Open-source tools for geospatial analysis.
- Digital spatial and mobility data.
- Preprocessing mobility data.
- Individual and collective mobility laws.
- Next-location prediction.
- Flow generation.
- Applications.

## Information retrieval (289AA) (6 ECTS)

*Semester:* 1

*Contact Person:* Prof. Rossano VENTURINI (<mailto:rossano.venturini@unipi.it>)

### *Objectives*

Study, design and analysis of IR systems which are efficient and effective to process, mine, search, cluster and classify documents, coming from textual as well as any unstructured domain. In the lectures, we will:

- study and analyze the main components of a modern search engine: Crawler, Parser, Compressor, Indexer, Query resolver, Query and Document annotator, Results Ranker;
- dig into some basic algorithmic techniques which are now ubiquitous in any IR application for data compression, indexing and sketching;
- describe few other IR tools which are used either as a component of a search engine or as independent tools and build up the previous algorithmic techniques, such as: Classification, Clustering, Recommendation, Random Sampling, Locality Sensitive Hashing.

### *Syllabus*

- Search engines.
- Crawling, Text analysis, Indexing, Ranking.
- Storage of Web pages and (hyper-)link graph.
- Results processing and visualization.
- Other data types: XML, textual DBs.
- Data processing for IR tools.
- Data streaming.
- Data sketching.
- Data compression.
- Data clustering (sketch).

---

## **Machine Learning (654AA) (9 ECTS)**

*Semester:* 1

*Contact Person:* Prof. Alessio MICHELI (<mailto:micheli@di.unipi.it>)

### *Objectives*

We introduce the principles and the critical analysis of the main paradigms for learning from data and their applications. The course provides the Machine Learning basis for both the aims of building new adaptive Intelligent Systems and powerful predictive models for intelligent data analysis.

### *Syllabus*

- Computational learning tasks for predictions, learning as function approximation, generalization concept.
- Linear models and Nearest-Neighbors (learning algorithms and properties, regularization).
- Neural Networks (MLP and deep models, SOM).
- Probabilistic graphical models.
- Principles of learning processes: elements of statistical learning theory, model validation.
- Support Vector Machines and kernel-based models.
- Introduction to applications and advanced models.
- Application project: implementation and use of ML/NN models with emphasis to the rigorous application of validation techniques.

## **Programmatic advertising (634AA) (6 ECTS)**

*Semester* 1

*Contact Person* Prof. Nicola CIARAMELLA (<mailto:ciaramella@noesis-research.com>)

### *Objectives*

The course aims at providing students with a conceptual framework and a toolbox for optimization of online advertising campaigns (inside sites, apps, games). At the end of the course the student should be able to design and possibly implement real-life systems for optimization of campaigns performance, intended in financial and marketing terms. The required mathematical background is limited to basic differential calculus and probability theory. The treatment is quantitative and concepts will be translated in formulas and algorithms. Nevertheless, focus will be on intuition and business meaning more than on formal rigor.

### *Syllabus*

- The online advertising ecosystem. Advertisers, publishers, business intermediaries, technology providers, data providers.
- Trends and Programmatic Advertising.
- Online advertising campaign management: design, targeting, creation, monitoring, optimization and reporting.
- Data about people and their behavior. Classical segmentation, micro-segmentation, one-to-one relationships.
- Data management platforms.
- The publisher problem. Basic micro-economic concepts and decision theory: expected utility, marginal utility, pricing, decision trees, value of information, risk and uncertainty, opportunity cost, equilibrium and optimality.
- The advertiser problem. Market segmentation, customer profiling.
- The advertisers-publishers game.
- Forecasting visitors and campaigns behavior. Classical methods: linear regression, logistic regression, time series analysis. Factorization methods. Markovian methods.
- Learning and optimization. Facing uncertainty. The Exp-Exp dilemma. Multi-armed bandits. Reinforcement learning.

---

## Social network analysis (668AA) (6 ECTS)

*Semester:* 2

*Contact Person:* Prof. Dino PEDRESCHI (<mailto:pedre@di.unipi.it>)

*Contact Person:* Prof. Giulio ROSSETTI (<mailto:giulio.rossetti@gmail.com>)

### *Objectives*

Over the past decade there has been a growing public fascination with the complex “connectedness” of modern society. This connectedness is found in many contexts: in the rapid growth of the Internet and the Web, in the ease with which global communication now takes place, and in the ability of news and information as well as epidemics and financial crises to spread around the world with surprising speed and intensity. These are phenomena that involve networks and the aggregate behavior of groups of people; they are based on the links that connect us and the ways in which each of our decisions can have subtle consequences for the outcomes of everyone else. This short course is an introduction to the analysis of complex networks, with a special focus on social networks and the Web: its structure and function, and how it can be exploited to search for information. Drawing on ideas from computing and information science, applied mathematics, economics and sociology, the course describes the emerging field of study that is growing at the interface of all these areas, addressing fundamental questions about how the social, economic, and technological worlds are connected.

### *Syllabus*

Graph theory and social networks:

- Graphs.
- Social, information, biological and technological networks.
- Strong and weak ties.
- Networks in their surrounding context.

The World Wide Web:

- The structure of the Web.
- Link analysis and Web search.
- Web mining and sponsored search markets.

Network dynamics:

- Information cascades.
- Power laws and rich-get-richer phenomena.
- The small-world phenomenon.
- Epidemics.

## **Text analytics (635AA) (6 ECTS)**

*Semester:* 1

*Contact Person:* Prof. Francesco SPINNATO (<mailto:francesco.spinnato@unipi.it>)

### *Objectives*

The course targets text analytics systems and advanced AI applications to address business problems by discovering and presenting knowledge locked in textual data. In addition to traditional text analytics, students will learn how to design and implement agentic systems, such as agentic Retrieval-Augmented Generation (RAG), which combine large language models with reasoning mechanisms and autonomous decision-making capabilities. The objective is to recognize situations where text analytics and agentic systems can solve information processing needs, identify the analytic task or process that best models the business problem, select appropriate resources, methods, and tools, and apply them to collected text data. Applications will include information extraction, sentiment analysis, spam and fake post detection, summarization, and advanced use cases such as knowledge-grounded conversational agents and workflow automation through agentic LLMs.

### *Syllabus*

- Disciplinary background: Natural Language Processing, Information Retrieval, Machine Learning, and Generative AI.
- Linguistic essentials: words, lemmas, morphology, PoS, syntax.
- Basic text processing: regular expressions, tokenization.
- Data gathering: Twitter API, scraping.
- Basic modeling: collocations, language models.
- Libraries and tools: NLTK, Keras, Hugging Face Transformers, LangChain.
- Agentic systems fundamentals:
  - Retrieval-Augmented Generation concepts and architectures.
  - Integration of LLMs with external knowledge sources.
  - Prompt engineering and grounding strategies.
  - Autonomous agents and tool orchestration.
- Applications:
  - Classification and clustering.
  - Sentiment analysis and opinion mining.
  - Information and relation extraction.
  - Entity linking.
  - Spam detection (mail, phishing, blogs, reviews).
  - Summarization and question answering with RAG.
  - Agent-based workflows for business intelligence.

---

## Visual analytics (602AA) (6 ECTS)

*Semester:* 2

*Contact Person:* Prof. Salvatore RINZIVILLO (<mailto:rinzivillo@isti.cnr.it>)

### *Objectives*

The availability of large data sources provides new opportunities for understanding patterns and behaviors of modern society. The information from these sources requires effective visualization methods to extract meaningful information from the data and facilitate the interpretation of very complex phenomena. The objective of the course is to present an overview of basic methods and visualization techniques for effective presentation of information from different sources: structured data (relational hierarchies, trees), relational data (social networks), temporal data, spatial data and data space-time. We will present and discuss several case study scenarios with the existing methods and tools.

### *Syllabus*

#### Visual Metaphors for Information

- Hierarchical and structured data.
- Relational and graph-based data.
- Temporal Data.
- Spatial data.
- Spatio-temporal data.
- Unstructured information (text).

#### Methods and Tools

- Overview of existing visual analytics environments.
- Visual Analytics Process.
- Definition of a Knowledge Discovery process.
- Framework for VA.
- Visual exploration and analytics of data.
- Case studies.

### A.3 Elective subjects from the GR2 group

#### Advanced Statistical Learning (0106A) (6 ECTS)

*Semester:* 2

*Contact Person (Module I):* Prof. Salvatore RUGGIERI (<mailto:salvatore.ruggieri@unipi.it>)

*Contact Person (Module II):* Prof.ssa Francesca CHIAROMONTE (<mailto:francesca.chiaromonte@santannapisa.it>)

#### *Objectives*

The course covers advanced statistical methods for machine learning. Module 1 introduces advanced statistical foundations for machine learning, focusing on robustness, decision-making under uncertainty, calibration, Bayesian and causal inference, and modern representation techniques. Module 2 addresses high dimensional and large-scale data challenges, emphasizing feature selection, dimension reduction, and strategies for managing massive or imbalanced datasets. Theoretical concepts are complemented by practical exercises and project work using the R or Python programming languages.

#### *Syllabus*

##### Module 1 (INF/01, 3 ECTS)

- Noise and robust statistics.
- Statistical decision theory.
- Classifier calibration, conformal methods and learning to reject/defer.
- Bayesian inference.
- Causal inference: structured causal model, potential outcome model.
- PCA and embeddings in LLM.

##### Module 2 (SECS-S/01, 3 ECTS)

- Feature selection and regularization techniques for high-dimensional Linear and Generalized Linear Models.
- Feature screening algorithms for ultra-high dimensional supervised problems.
- Supervised dimension reduction; SuXicient Dimension Reduction and related techniques.
- Subsampling/partitioning approaches for ultra-high sample sizes.
- Under- and oversampling approaches for data rebalancing.



---

## **Financial Analysis and Performance Measurement (0003P) (9 ECTS)**

*Semester:* 1

*Contact Person:* Prof. Mariarita PIEROTTI (<mailto:mariarita.pierotti@unipi.it>)

### *Objectives*

This course is designed to enhance the skills of students with an economic background. It particularly focuses on the use of financial statements for analysis and valuation. Students will witness accounting principles being applied in real-life scenarios, specifically in the realms of financial analysis and equity valuation. Throughout this course, participants will learn how to utilize accounting and financial information to analyze corporate financial performance and conduct equity valuation. The course will also introduce techniques for analyzing financial statements for the purpose of financial analysis and evaluation, such as for acquisitions, restructurings, investments, and for assessing the value generated by different strategic scenarios. In the second part of the course, support will be provided for business decisions, evaluating organizational performance, and making assessments using both financial and non-financial information.

### *Syllabus*

The first part of the course aims to enhance your understanding of financial statement analysis in order to assess performance. It begins with a brief overview of financial analysis and its considerations, followed by the following steps:

- Understanding balance sheets, income statements, and statements of cash flow.
- Quantitative analysis.
- Transitioning from quantitative to qualitative analysis.
- Identifying red flags in financial analysis.

The second part of the course focuses on management and control systems to ensure effective enterprise control. Utilizing financial controls requires managers to make decisions regarding:

- Responsibility structures.
- Performance measures.
- Performance evaluations.
- Rewards and incentives.

## **Fundamentals of business management (627PP) (9 ECTS)**

*Semester:* 1

*Contact Person:* Prof. Mariarita PIEROTTI (<mailto:mariarita.pierotti@unipi.it>)

### *Objectives*

The course is designed to introduce students from different backgrounds (especially from STEM disciplines) to the most relevant topics and concepts characterizing business management, and to provide them with the ability to apply business management knowledge to practical tasks. The topics proposed are relevant to a variety of application sectors, and cover both theoretical concepts and relevant managerial implications. The course will also introduce students to the basic concepts of financial and managerial accounting, including concepts of costs, volumes and profits, and their relationships, budgeting, performance measurement and evaluation. Students will increase their familiarity with the accounting process, and will learn to read and understand financial statements. The course will also discuss the economics and management of strategic decisions and human capital.

### *Syllabus*

- The firm and the market.
- What is a firm?
- Theories of the firm.
- The firm as a sustainable system.
- Where does a firm work, interact and evolve?
- Value creation and firm's objectives.
- The competitive advantage and the entry strategies.
- The firm's functions.
- Production, Research & Development, Accounting, Finance, Innovation.
- Processes and techniques of marketing.
- Operations management.
- Procurement and logistics.
- Finance.
- Principles of financial accounting and budget.
- Firm's value, finance and capital structure.
- Resources management, strategy, business models.
- Economics and management of strategic decisions.
- Governance, networking and strategic collaboration.
- Human resources management practices.

*Notice:* this course cannot be chosen if the student bachelor is from the *Economics* area.

---

## Management practice (629PP) (6 ECTS)

*Semester:* 2

*Contact Person:* Prof. Giulio FERRIGNO (<mailto:giulio.ferrigno@santannapisa.it>)

### *Objectives*

This advanced course aims at providing students with knowledge of the main topics and management practices characterizing today's competitive environment. A special emphasis will be put on the role of enabling technologies and on management practices in innovative firms. The course will also introduce students to entrepreneurial practices in information science (how to develop and bring to the market new products and services based on embedded systems and high-tech solutions).

### *Syllabus*

#### 1) The innovative firm: theoretical concepts and management implications

- Knowledge, invention and innovation.
- Technological paradigms: nature and evolution.
- Sources of innovation.
- Innovation typologies and dynamics.
- Innovation diffusion and market barriers.
- The innovative firm: resources, competencies and boundaries.
- Quantitative indicators of innovative activities.
- Innovation and firm growth.
- The geography of R&D: knowledge and innovation.
- Entrepreneurial practices in information science.

#### 2) Management practices and enabling technologies

- Technological paradigms in information science.
- Architecture machine (r)evolution and firms' organization.
- Software (r)evolution and firms' key competencies evolution.
- Network (r)evolution and firms' boundaries.
- AI (r)evolution and Industry 4.0.

**Strategic and competitive intelligence (787II) (6 ECTS)***Semester:* 2*Contact Person:* Prof. Irene SPADA (<mailto:irene.spada@unipi.it>)*Objectives*

Competitive Intelligence (CI) activities have goals such as proactively detecting opportunities or threats, eliminating or reducing blind-spots, risks and/or surprises; and reducing reaction time to competitor and marketplace changes. The final product of any worthwhile CI activity is what practitioners' term actionable intelligence, i.e. intelligence that management can act upon. It is more than analysing competitors: it is a process for gathering information, converting it into intelligence (about products, customers, competitors, and any aspect of the environment) and then using it in decision making. In this sense, big data brings big change to CI. The course aims at developing CI projects based on the application on Data Science tools and methodologies. The course includes in-class seminars that introduce the fundamentals of competitive intelligence, including systems and strategic thinking. It provides many Data Science tools and techniques. Students will apply these tools in groups when analysing a preselected case company. They are expected to present early-stage versions of their reports, and, in the final workshop, they will present the results of their CI analysis, which is then discussed in plenary.

*Syllabus*

- [1] SCI FUNDAMENTALS: VUCA, Ansoff Model, surprise in business, risk & uncertainty, applications of CI, CI cycle
- [2] DATA INTELLIGENCE BASICS: Data Cycle for CI (Plan-Gather-Analyse-Interpret), Data source for CI (technical, scientific and market source).
- [3] DATA SCIENCE FOR SCI PROJECTS: (1) Basics: text analysis; (2) Advanced: Named Entity Recognition (NER), topic modelling, Social Network Analysis (SNA) and data visualization.
- [4] DATA SCIENCE PROJECT DESIGN: Scoping, Key Intelligence Topics (KITs) and Key Intelligence Questions (KIQs), metrics, management, result, visualization.
- [5] SCI APPLICATION LAB: How to extract intelligence from scientific papers. How to extract intelligence from IP. How to extract intelligence from HR and other sources (e.g., Wikipedia, Social Media, Online Reviews).

---

## A.4 Elective subjects from the GR3 group

### Legal issues in data science (381NN) (6 ECTS)

*Semester:* 2

*Contact Person:* Prof. Giovanni COMANDÈ (<mailto:giovanni.comande@sssup.it>)

*Contact Person:* Prof.ssa Denise AMRAM (<mailto:denise.amram@gmail.com>)

#### *Objectives*

The digital economy and the digital society harness the power of big data, computational capacity, innovation and interconnection. Every human activity is mediated by information technologies. Today's technologies enable unprecedented exploitation of information, being it small or big data, for any thinkable purpose, but mostly in business and surveillance with the ensuing legal and ethical anxieties and constraints. Algorithms are regularly used for mining data, offering unexplored patterns and deep non-causal analyses to those businesses able to exploit these advances. Yet, these innovations need to be properly framed in the existing legal background, fit in the existing set of guarantees of fundamental rights and freedoms, coherently policy related to reap the richness of big and open data and administration while empowering equally all players. For these aims data protection plays a significant role. The course aims at enabling students to work on algorithms and data mining techniques in ways that are compliant to the applicable legal framework and aware of the interplay between techniques and normative rules.

#### *Syllabus*

The Algorithmic Society: the Classifying Society:

- Background and Overview, Surveillance Society.
- Big Other, Networks of Control, Predicting Behavior.
- People Analytics, Behavioural “Nudging”.
- New Emerging Human Rights in the age of Behavioral Data Science and Neurotechnologies.
- Towards “Mental Privacy” and “Decision Integrity”.
- Legal and ethical implication of computational capacity.

Building Legally-Compliant Algorithms:

- Legal Pitfalls of Algorithms, The Problems of Personalization, Data Handling & Sharing.
- Deploying Algorithms for Human Rights: Complications & Challenges.
- Classification of Algorithms in the Information Society.
- Legal Implications and Business Applications, Exploitation of Public Sector Data.
- Competition Law in the Age of Algorithms, Transparency.
- Accountability and traceability of algorithm based decision-making.
- Accountability in the Machine Learning Context.
- Technical and Legal Options to Enhance Transparency & Accountability.
- Legal Liability for Algorithm Autocomplete (ISP Liability).
- Open Data Governance, Data Ethics.

General principles of privacy law: The American approach, The European approach.

The General Data Protection Regulation:

- Notions and principles, GDPR global reach and compliance.
- Google Spain Decision.
- Invalidation of Data Retention Directive (US Safe Harbour Decision)/Schrems.

Privacy in operation:

- Privacy-by-Design, GDPR Solutions: The Right to an Explanation, etc.
- Notions of Privacy in the Algorithmic Age, Privacy from the Government.
- Surveillance Capitalism, Governance by Proxy, Privacy from Private Entities.
- Privacy from Platforms, Privacy from Employers, Privacy from our Devices (IoT).

Comparative Perspectives & Crossborder Issues:

- Comparative Privacy and security Regimes: GDPR vs. USA.
- Comparative Privacy and security Regimes: GDPR vs. dChina.

**Model-Driven Decision-Making Methods (666AA) (6 ECTS)**

*Semester:* 1

*Contact Person:* Prof. Antonio FRANGIONI (<mailto:antonio.frangioni@unipi.it>)

*Objectives*

The course will enable the student to use software tools based on mathematical optimization techniques for the support to complex decisions (mainly at the corporate/industrial level), focusing on practical aspects of these tools. The main aim is to familiarize the students with the specific computer science aspects of these activities, such as data preparation and validation, the development of complex mathematical models, the knowledgeable use of the corresponding solution algorithms, the impact on this process of data uncertainty and the available methodologies to tackle this problem.

*Syllabus*

- Decision theory, decision processes.
- Architecture of decision support systems.
- Applications-oriented overview of Linear Programming and Integer Linear Programming problems.
- Algorithms and solvers of Linear Programming and Integer Linear Programming problems.
- Methodologies for improving performances of the algorithms.
- Data uncertainty issues within optimization methods.

---

## **Programming and Algorithms for Data Science (0104A) (12 ECTS)**

*Semester:* 1

*Contact Person:* Prof. Salvatore TRANI (<mailto:salvatore.trani@isti.cnr.it>)

*Contact Person:* Prof. Rossano VENTURINI (<mailto:rossano.venturini@unipi.it>)

### *Objectives*

This course provides the foundations of programming and algorithmic thinking for Data Science students. It combines fundamental programming skills in Python with the study of classical algorithms and data structures, focusing on their use for solving data processing and analysis problems. An intensive laboratory activity will complement the lectures. Students will learn how to design efficient solutions, analyze computational complexity, and implement them effectively using modern Python tools and libraries. The course comprises lectures and practical sessions in computer labs.

### *Syllabus*

#### Module 1: Laboratory of Programming

- Python development tools.
- Python data types and data structures.
- Variable, functions, and conditional statements.
- Iterative vs recursive computation.
- Imperative programming vs object-oriented programming.
- Problem-solving and basic algorithms.

#### Module 2: Algorithms and Data Structures

- Introduction and basic definitions: algorithm, problem, instance.
- Computational complexity analysis of algorithms.
- Sorting: Mergesort, Quicksort, and Heapsort.
- Searching: Binary Search, Binary Search Tree, Trie, and Hashing.
- Algorithms on Trees: representation and traversals.
- Algorithms on Graphs: representation, traversals, and most important problems.

## A.5 Elective subjects from the Table 2.4 group

The subjects from Table 2.4 of the *Informatics* area are offered by the Master Degree program in Computer Science. Detailed information will be available at the website:

<https://www.di.unipi.it/en/education/mcs>

**Important notice:** the timetable of these subjects will not be included in the official timetable of the Data Science and Business Informatics programme. Please, check the website above for timetables.



---

## Corsi in Italiano

**Ingegneria del software\* (271AA) (6 ECTS)**

*Title in English:* Software Engineering

*Semester:* 2

*Contact Person:* Prof.ssa Laura SEMINI (<mailto:laura.semini@unipi.it>)

*Obiettivi*

Fornire le metodologie e strumenti per la progettazione, realizzazione, verifica, validazione e misurazione di sistemi software.

*Syllabus*

- Processo di sviluppo software: problemi della produzione del software, modelli di ciclo di vita.
- Analisi del dominio: modelli statici (classi e associazioni) e dinamici (attività, macchine a stati).
- Analisi dei requisiti: modello statico (casi d'uso) e dinamici (narrative, diagrammi di robustezza).
- Progettazione architetturale: modelli statici (viste strutturali e logistiche) e dinamici (vista componenti/connettori).
- Progettazione di dettaglio: modello statico delle componenti (strutture composite) e modello dinamico (interazioni).
- Verifiche e prove: obiettivi e pianificazione delle verifiche, progettazione e valutazione delle prove.