



SAPIENZA  
UNIVERSITÀ DI ROMA

# Data Science (2024)

## Il corso

Codice corso: 32344

Classe di laurea: LM-Data

Durata: 2 anni

Lingua: ENG

Modalità di erogazione:

Dipartimento: INGEGNERIA INFORMATICA, AUTOMATICA E GESTIONALE "ANTONIO RUBERTI"

## Presentazione

Il corso di laurea magistrale in Data Science si caratterizza per un'offerta didattica interdisciplinare che raccoglie contributi dell'ingegneria, dell'informatica, della statistica, delle scienze economiche e organizzative, insieme a conoscenze specifiche dei principali domini applicativi di Data Science. In particolare, la laurea magistrale in Data Science offre le conoscenze professionali adeguate per lo sviluppo delle tecnologie di raccolta, gestione, elaborazione e analisi dei big data, e la conseguente traduzione in informazioni fondamentali per il processo conoscitivo e decisionale all'interno dei settori innovativi di business e sociali. Il corso di studio è di durata biennale e prevede un primo insieme di 39 CFU su settori scientifico disciplinari caratterizzanti miranti a fornire le conoscenze statistiche, ingegneristiche e informatiche di base necessarie per lo sviluppo degli strumenti software e delle infrastrutture necessarie per la raccolta, l'elaborazione, e l'organizzazione delle grandi moli di dati e dei modelli matematico statistici utili per la loro analisi. Nei 39 CFU sono inclusi almeno 10 CFU di attività di laboratorio. Tali corsi caratterizzanti saranno obbligatori per tutti gli studenti. I 39 CFU obbligatori si dividono in 27 CFU sulle tecnologie informatiche e 12 CFU nelle discipline statistiche. Lo studente potrà quindi scegliere fino a 30 CFU di indirizzo su settori scientifico disciplinari caratterizzanti. Almeno 6 dovranno essere scelti tra le discipline umane, sociali, giuridiche ed economiche. Tali insegnamenti sono volti alla formazione di un profilo professionale che unisce le competenze ingegneristiche e informatiche con le competenze statistiche e gestionali, economiche e giuridiche. Tali competenze devono essere sviluppate insieme ad una profonda conoscenza del contesto economico, sociale e organizzativo in cui le metodologie di Data Science si vanno ad applicare. Sono previsti inoltre 3 CFU per Altre Attività Formative che includono attività di tirocinio presso aziende e partecipazione a Training Camps tematici. Il percorso formativo si completerà con 12 CFU di attività su settori scientifici disciplinari affini e 12 CFU a scelta libera. Non sono previsti obblighi di frequenza se non per attività di laboratorio e attività pratiche. Tutti gli insegnamenti sono erogati in lingua inglese. Le conoscenze raggiunte sono valutate tramite prove intermedie, discussione di lavori di gruppo o elaborati redatti singolarmente dai discenti e accertate tramite esami di tipo tradizionale. Il percorso formativo consente al laureato magistrale in Data Science di trovare occupazione presso piccole e medie imprese, grandi aziende, pubblica amministrazione, amministrazioni locali, istituti di ricerca pubblici e privati ed enti noprofit. Il Dottorato di Ricerca e i Master di II livello sono altre possibilità a cui lo studente in Data Science potrebbe rivolgersi al termine del suo percorso di studio.

# Percorso formativo

Curriculum unico

## 1° anno

Insegnamento	Semestre	CFU	Lingua
1047221   ALGORITHMIC METHODS OF DATA MINING AND LABORATORY	1°	9	ENG

### Obiettivi formativi

? The course presents the main algorithmic techniques of data mining, necessary for data science. They offer to the student the basis for analyzing data for a variety of applications that deal with semistructured or unstructured data, such as textual data, transactions, and graph and information-network data. At the end of the course the student will have a knowledge of the main theoretical ideas of data mining, as well as some basic knowledge and experience in using programming tools for analyzing and mining data.

1047264   FUNDAMENTALS OF DATA SCIENCE AND LABORATORY	1°	9	ENG
---	----	---	-----

### Obiettivi formativi

Learning from data in order to gain useful predictions and insights. At the end of the course students will have an understanding of the basic programming skills needed for data analysis and visualization. They will also have familiarity of the typical data processing workflow of data preparation and scraping, visualization and exploratory analysis and final statistical modeling. Students will become familiar with the main Python libraries for data science.

10589600   STATISTICAL METHODS IN DATA SCIENCE AND LABORATORY	1°	12	ENG
---	----	----	-----

**Obiettivi formativi**

## Learning goals

Statistical Methods in Data Science is a two-semester course aimed at providing the fundamental tools for:

setting up probabilistic models;

understanding the basic principles of the main inferential problems: estimation, hypothesis testing, model checking and forecasting;

understanding and contrasting the two main inferential paradigms, namely frequentist and Bayesian statistics;

implementing inference on observed data through both optimization and simulation-based (approximation) techniques such as:

Bootstrap

Monte Carlo

Monte Carlo Markov Chain (MCMC)

understanding comparative merits of alternative strategies

developing statistical computations within a suitable software environment like R ([www.r-project.org](http://www.r-project.org)), OpenBUGS (

<http://openbugs.net/w/FrontPage>) and STAN (<http://mc-stan.org/>).

## Knowledge and understanding

On successful completion of this course, students will:

know the main statistical principles, inferential problems, paradigms and algorithms;

assess the empirical and theoretical performance of different modeling approaches;

know the main platforms, programming languages to develop effective implementations.

## Applying knowledge and understanding

Besides the understanding of theoretical aspects, thanks to applied homeworks and a dedicated laboratory in the second semester focused on Bayesian modeling, students will be constantly challenged to use and evaluate all the techniques they have learned as well as to propose new modelization suitable for specific tasks at hand.

## Making judgements

On successful completion of this course, students will develop a positive critical attitude towards the empirical and theoretical evaluation of statistical methodologies and results.

## Communication skills

In preparing the report and oral presentation for the final project of the second semester laboratory, students will learn how to effectively communicate information, ideas, problems and solutions to specialists but also to a general audience.

## Learning skills

In this course the students will develop the skills necessary for a successful understanding and application of new statistical methodologies together with their effective implementation. The goal is of course to grow an active attitude towards continued learning throughout a professional career.

STATISTICAL METHODS  
IN DATA SCIENCE AND  
LABORATORY I

1°

9

ENG

**Obiettivi formativi**

## Learning goals

Statistical Methods in Data Science is a two-semester course aimed at providing the fundamental tools for:

setting up probabilistic models;  
 understanding the basic principles of the main inferential problems: estimation, hypothesis testing, model checking and forecasting;  
 understanding and contrasting the two main inferential paradigms, namely frequentist and Bayesian statistics;  
 implementing inference on observed data through both optimization and simulation-based (approximation) techniques such as:  
 Bootstrap  
 Monte Carlo  
 Monte Carlo Markov Chain (MCMC)  
 understanding comparative merits of alternative strategies  
 developing statistical computations within a suitable software environment like R ([www.r-project.org](http://www.r-project.org)), OpenBUGS (<http://openbugs.net/w/FrontPage>) and STAN (<http://mc-stan.org/>).

## Knowledge and understanding

On successful completion of this course, students will:

know the main statistical principles, inferential problems, paradigms and algorithms;  
 assess the empirical and theoretical performance of different modeling approaches;  
 know the main platforms, programming languages to develop effective implementations.

## Applying knowledge and understanding

Besides the understanding of theoretical aspects, thanks to applied homeworks and a dedicated laboratory in the second semester focused on Bayesian modeling, students will be constantly challenged to use and evaluate all the techniques they have learned as well as to propose new modelization suitable for specific tasks at hand.

## Making judgements

On successful completion of this course, students will develop a positive critical attitude towards the empirical and theoretical evaluation of statistical methodologies and results.

## Communication skills

In preparing the report and oral presentation for the final project of the second semester laboratory, students will learn how to effectively communicate information, ideas, problems and solutions to specialists but also to a general audience.

## Learning skills

In this course the students will develop the skills necessary for a successful understanding and application of new statistical methodologies together with their effective implementation. The goal is of course to grow an active attitude towards continued learning throughout a professional career.

**Insegnamento****Semestre****CFU****Lingua****Obiettivi formativi**

## Obiettivi generali:

Lo scopo del corso è fornire agli studenti gli strumenti per comprendere i principi del networking e le principali tecnologie di rete. Il corso è focalizzato sull'evoluzione della rete Internet per il supporto dei big data e del cloud, con particolare attenzione alle soluzioni di rete per i data centers. La prima parte del corso sarà necessaria per rendere omogeneo il livello della classe e per definire i concetti e i termini tecnici di base. Il corso prevede anche l'utilizzo di un emulatore di rete e di un analizzatore di traffico per lo svolgimento di attività pratiche di laboratorio.

## Obiettivi specifici:

Conoscenza e capacità di comprensione: conoscere i principali protocolli di rete per la realizzazione di una rete IP.

Conoscenza e capacità di comprensione applicate: saper applicare i principi del networking per realizzare una rete emulata funzionante e per analizzare in maniera critica il traffico all'interno di una rete

Autonomia di giudizio: capacità di individuare criticamente gli elementi di debolezza delle soluzioni architetturali studiate nello scenario di un data center per la gestione dei big data

Capacità di apprendere: capacità di proseguire gli studi successivi riguardanti tematiche avanzate di networking.

10589600 |  
STATISTICAL METHODS  
IN DATA SCIENCE AND  
LABORATORY

2°

12

ENG

**Obiettivi formativi**

## Learning goals

Statistical Methods in Data Science is a two-semester course aimed at providing the fundamental tools for:

setting up probabilistic models;  
 understanding the basic principles of the main inferential problems: estimation, hypothesis testing, model checking and forecasting;  
 understanding and contrasting the two main inferential paradigms, namely frequentist and Bayesian statistics;  
 implementing inference on observed data through both optimization and simulation-based (approximation) techniques such as:  
 Bootstrap  
 Monte Carlo  
 Monte Carlo Markov Chain (MCMC)  
 understanding comparative merits of alternative strategies  
 developing statistical computations within a suitable software environment like R ([www.r-project.org](http://www.r-project.org)), OpenBUGS (<http://openbugs.net/w/FrontPage>) and STAN (<http://mc-stan.org/>).

## Knowledge and understanding

On successful completion of this course, students will:

know the main statistical principles, inferential problems, paradigms and algorithms;  
 assess the empirical and theoretical performance of different modeling approaches;  
 know the main platforms, programming languages to develop effective implementations.

## Applying knowledge and understanding

Besides the understanding of theoretical aspects, thanks to applied homeworks and a dedicated laboratory in the second semester focused on Bayesian modeling, students will be constantly challenged to use and evaluate all the techniques they have learned as well as to propose new modelization suitable for specific tasks at hand.

## Making judgements

On successful completion of this course, students will develop a positive critical attitude towards the empirical and theoretical evaluation of statistical methodologies and results.

## Communication skills

In preparing the report and oral presentation for the final project of the second semester laboratory, students will learn how to effectively communicate information, ideas, problems and solutions to specialists but also to a general audience.

## Learning skills

In this course the students will develop the skills necessary for a successful understanding and application of new statistical methodologies together with their effective implementation. The goal is of course to grow an active attitude towards continued learning throughout a professional career.

STATISTICAL METHODS  
 IN DATA SCIENCE AND  
 LABORATORY II

2°

3

ENG

**Obiettivi formativi**

## Learning goals

Statistical Methods in Data Science is a two-semester course aimed at providing the fundamental tools for:

setting up probabilistic models;  
 understanding the basic principles of the main inferential problems: estimation, hypothesis testing, model checking and forecasting;  
 understanding and contrasting the two main inferential paradigms, namely frequentist and Bayesian statistics;  
 implementing inference on observed data through both optimization and simulation-based (approximation) techniques such as:  
 Bootstrap  
 Monte Carlo  
 Monte Carlo Markov Chain (MCMC)  
 understanding comparative merits of alternative strategies  
 developing statistical computations within a suitable software environment like R ([www.r-project.org](http://www.r-project.org)), OpenBUGS (<http://openbugs.net/w/FrontPage>) and STAN (<http://mc-stan.org/>).

## Knowledge and understanding

On successful completion of this course, students will:

know the main statistical principles, inferential problems, paradigms and algorithms;  
 assess the empirical and theoretical performance of different modeling approaches;  
 know the main platforms, programming languages to develop effective implementations.

## Applying knowledge and understanding

Besides the understanding of theoretical aspects, thanks to applied homeworks and a dedicated laboratory in the second semester focused on Bayesian modeling, students will be constantly challenged to use and evaluate all the techniques they have learned as well as to propose new modelization suitable for specific tasks at hand.

## Making judgements

On successful completion of this course, students will develop a positive critical attitude towards the empirical and theoretical evaluation of statistical methodologies and results.

## Communication skills

In preparing the report and oral presentation for the final project of the second semester laboratory, students will learn how to effectively communicate information, ideas, problems and solutions to specialists but also to a general audience.

## Learning skills

In this course the students will develop the skills necessary for a successful understanding and application of new statistical methodologies together with their effective implementation. The goal is of course to grow an active attitude towards continued learning throughout a professional career.

A SCELTA DELLO  
STUDENTE

2°

6

ENG

**Obiettivi formativi**

Among other training activities are provided 12 credits are chosen by the student.

Gruppo opzionale A

Gruppo opzionale B

Gruppo opzionale C

**2° anno**

<b>Insegnamento</b>	<b>Semestre</b>	<b>CFU</b>	<b>Lingua</b>
A SCELTA DELLO STUDENTE	1°	6	ITA
<b>Obiettivi formativi</b>			
Among other training activities are provided 12 credits are chosen by the student.			
AAF1149   altre conoscenze utili per l'inserimento nel mondo del lavoro	2°	3	ITA
<b>Obiettivi formativi</b>			
The specific aim is to enable the student to assist him with the more specific knowledge for inclusion in the future world of work.			
AAF1022   PROVA FINALE	2°	24	ENG
<b>Obiettivi formativi</b>			
The student will present and discuss the results of a technical activity, producing a written thesis supervised by a professor and showing the ability to master the methodologies of data science			
Gruppo opzionale B Gruppo opzionale D			

### Gruppi opzionali

Lo studente deve acquisire 6 CFU fra i seguenti esami

<b>Insegnamento</b>	<b>Anno</b>	<b>Semestre</b>	<b>CFU</b>	<b>Lingua</b>
1047208   STATISTICAL LEARNING	1°	2°	6	ENG

**Insegnamento****Anno****Semestre****CFU****Lingua****Obiettivi formativi**

Devising new machine learning methods and statistical models is a fun and extremely fruitful “art”. But these powerful tools are not useful unless we understand when they work, and when they fail. The main goal of statistical learning theory is thus to study, in a

statistical framework, the properties of learning algorithms mainly in the form of so-called error bounds.

This course introduces the techniques that are used to obtain such results, combining methodology with theoretical foundations and computational aspects. It treats both the basic principles to design successful learning algorithms and the “science” of analyzing an algorithm’s statistical properties and performance guarantees.

Theorems are presented together with practical aspects of methodology and intuition to help students develop tools for selecting appropriate methods and approaches to problems in their own data analyses.

Methods for a wide variety of applied problems will be explored and implemented on open-source software like R ([www.r-project.org](http://www.r-project.org)), Keras (<https://keras.io/>) and TensorFlow (<https://www.tensorflow.org/>).

Knowledge and understanding

On successful completion of this course, students will:

know the main learning methodologies and paradigms with their strengths and weakness;

be able to identify a proper learning model for a given problem;

assess the empirical and theoretical performance of different learning models;

know the main platforms, programming languages and solutions to develop effective implementations.

Applying knowledge and understanding

Besides the understanding of theoretical aspects, thanks to applied homeworks and a final project possibly linked to hackathons or other data analysis competitions, the students will constantly be challenged to use and evaluate modern learning techniques and algorithms.

Making judgements

On successful completion of this course, students will develop a positive critical attitude towards the empirical and theoretical evaluation of statistical learning paradigms and techniques.

Communication skills

In preparing the report and oral presentation for the final project, students will learn how to effectively communicate original ideas, experimental results and the principles behind advanced data analytic techniques in written and oral form. They will also understand how to offer constructive critiques on the presentations of their peers.

Learning skills

In this course the students will develop the skills necessary for a successful understanding as well as development of new learning methodologies together with their effective implementation. The goal is of course to grow a active attitude towards continued learning throughout a professional career.

10606725 |  
OPTIMIZATION  
METHODS FOR  
DATA SCIENCE

1<sup>o</sup>2<sup>o</sup>

6

ENG

**Obiettivi formativi**

Lo scopo del corso è quello di introdurre gli studenti all'applicazione delle tecniche di ottimizzazione per l'addestramento in problemi di apprendimento

automatico. Si prevede che gli studenti acquisiscano competenze sui modelli standard usati in apprendimento automatico (Deep Networks and Support Vector

Machines), capiscano quale modello sia più opportuno utilizzare in ogni contesto, e vengano a conoscenza dei più recenti algoritmi di ottimizzazione per determinare i parametri (addestrare) di tali modelli che meglio si adattano ai dati disponibili.

10615930 |  
STOCHASTIC  
PROCESSES FOR  
DATA SCIENCE

1<sup>o</sup>2<sup>o</sup>

6

ENG

Insegnamento	Anno	Semestre	CFU	Lingua
<b>Obiettivi formativi</b>				
DESCRIZIONE GENERALE				
L' obiettivo di questo corso è di fornire una panoramica sui processi stocastici, tenendo presenti le applicazioni alla scienza dei dati. I processi stocastici e la probabilità sono importanti per la scienza dei dati in quanto possono essere usati per l'analisi e lo sviluppo di modelli concernenti un'ampia gamma di dati, dai dati finanziari ai dati provenienti da sensori. Il corso è diviso in tre parti: un'introduzione sui processi stocastici combinatori, una seconda parte sui processi gaussiani e una terza parte sulla causalità probabilistica. La programmazione in R, Matlab o Python è utile per il corso, ma non è essenziale. Negli esempi si useranno programmi in R.				
OBIETTIVI SPECIFICI:				
1. Conoscenza e comprensione: comprendere i concetti di base sui processi stocastici combinatori e sui processi gaussiani e le loro applicazioni alla scienza dei dati. Comprendere i fondamenti della causalità probabilistica e saperla applicare a problemi reali.				
2. Applicazioni: applicare processi stocastici a dati reali, usando linguaggi di programmazione come R, Matlab o Python.				
3. Autonomia di giudizio: analizzare i vantaggi e i limiti dei diversi processi stocastici e determinare il miglior modello da utilizzare per un determinato problema.				
4. Comunicazione: comunicare in modo efficace i processi stocastici, compresi vincoli progettuali, soluzioni e possibili applicazioni.				
5. Abilità di apprendimento: saper sviluppare studi nell' ambito dei processi stocastici per la scienza dei dati, inclusa la capacità di intraprendere ricerche in questo campo.				

Lo studente deve acquisire 18 CFU fra i seguenti esami

Insegnamento	Anno	Semestre	CFU	Lingua
1047205   CLOUD COMPUTING	1°	2°	6	ENG

Insegnamento	Anno	Semestre	CFU	Lingua
<b>Obiettivi formativi</b>				
Obiettivi generali:				
<p>Lo scopo del corso è quello di fornire agli studenti i concetti di base dei sistemi distribuiti e quindi di concentrarsi sulle tecnologie di cloud computing. Il corso copre aspetti teorici e pratici con un focus su esempi reali. Alla fine del corso si suppone che gli studenti siano in grado di scegliere, configurare e utilizzare i servizi cloud e progettare e distribuire architetture scalabili ed applicazioni elastiche.</p>				
Obiettivi specifici:				
<p>Conoscenza e capacità di comprensione:</p> <p>Al termine del corso, lo studente sarà in grado di descrivere e spiegare</p> <ul style="list-style-type: none"> <li>- i concetti generali dei sistemi distribuiti</li> <li>- il concetto di virtualizzazione di sistema e applicazione</li> <li>- i meccanismi e gli algoritmi utilizzati nel cloud computing</li> <li>- le tecnologie per lo storage cloud</li> <li>- i framework per l'elaborazione dei bigdata</li> <li>- i problemi di sicurezza informatica e le soluzioni nel cloud computing</li> </ul>				
<p>Conoscenza e capacità di comprensione applicate:</p> <p>Al termine del corso, lo studente sarà in grado di</p> <ul style="list-style-type: none"> <li>- progettare e implementare un'architettura scalabile e distribuire un'applicazione elastica</li> <li>- presentare risultati pratici sotto forma di rapporto tecnico</li> <li>- analizzare e presentare lavori scientifici</li> <li>- selezionare, configurare ed eseguire servizi cloud utilizzando la GUI e l'API di gestione offerte dai provider IaaS</li> <li>- progettare e configurare infrastrutture scalabili e applicazioni distribuite elastiche.</li> <li>- fare scelte progettuali che tengano conto dei problemi di sicurezza informatica</li> </ul>				
<p>Autonomia di giudizio:</p> <p>Al termine del corso, lo studente:</p> <ul style="list-style-type: none"> <li>- sarà in grado di valutare e confrontare le tecnologie cloud e i servizi cloud, nonché i framework di elaborazione dei big data</li> <li>- sarà in grado di identificare, valutare e confrontare soluzioni all'avanguardia</li> <li>- rafforzerà la sua capacità di pensiero critico</li> </ul>				
<p>Abilità comunicative:</p> <p>Al termine del corso lo studente:</p> <ul style="list-style-type: none"> <li>- sarà in grado di discutere e trasmettere la propria opinione sulle tecnologie cloud</li> <li>- sarà in grado di presentare l'analisi di un argomento selezionato a un vasto pubblico</li> </ul>				
<p>Capacità di apprendere:</p> <p>Durante il corso, lo studente svilupperà e migliorerà la sua capacità di pensiero critico attraverso lo studio e l'analisi di lavori scientifici e di documentazione tecnica. Inoltre, lo studente migliorerà la sua capacità di integrare informazioni da diverse fonti, ad es. libri, documenti tecnici/scientifici ed esperienze pratiche.</p>				
1047197   DATA MANAGEMENT FOR DATA SCIENCE	1°	2°	6	ENG

Insegnamento	Anno	Semestre	CFU	Lingua
<b>Obiettivi formativi</b>				
<p>The main goal of the course is to present the basic concepts of data management systems. The first part of the course introduces the main aspects of relational database systems, including basic functionalities, file and index organizations, and query processing. The second part of the course aims at presenting the main non-relational approaches to data management, in particular, multidimensional data management, large-scale data management, and open data management.</p>				
10606654   ADVANCED DATA MINING AND LANGUAGE TECHNOLOGY	1°	2°	6	ENG
<b>Obiettivi formativi</b>				
<p>The course will present fundamental technologies for advanced data mining applications. The course will start with presenting the methodologies for storing and retrieving information on the Web, mining application logs, mining social media, collaborative filtering and personalization. The course will also present the basic technologies for classification and learning, with emphasis on textual data sources. Applications will include mining of consumer preferences, online marketplaces, digital marketing and Natural Language Processing applications such as sentiment analysis. As part of the course students will carry on a field study on a relevant use case for a selected application.</p>				
10589621   ADVANCED MACHINE LEARNING	2°	1°	6	ENG

Insegnamento	Anno	Semestre	CFU	Lingua
<b>Obiettivi formativi</b>				
<p>General objectives:  The course will present to students advanced and most recent concepts of machine learning and their application in computer vision via deep neural network (DNN) models. It will include theory and practical coding, as well as a final hands-on project. Towards the coding assignments and the final project, the students will work in teams and present their ideas and project outcome to the class.</p> <p>Specific objectives  The first part of the course includes delving into state-of-the-art DNN models for classification and regression applied to detection (where the objects are in the image), pose estimation (whether people stand, sit or crunch) and re-identification (estimating a unique vector representation for each person). The course further discusses DNNs for multi-task objectives (joint detection, pose estimation, re-identification, segmentation, depth estimation etc). This first part would include DNNs which apply to video sequences, by leveraging memory (e.g. LSTMs) or attention (Transformers).  The second part of the course delves into models, training techniques and data manipulation for generalization, domain adaptation and meta-learning. Further to transfer learning (how pre-trained models may be deployed for other tasks), it discusses multi-modal (with different sensor modalities such as depth or thermal cameras) and self-supervision (e.g. training the DNN model by solving jigsaw puzzles) to auto-annotate large amounts of data. Finally, it presents domain adaptation (e.g. apply daytime-detectors for night vision) and meta-learning, a most recent framework to learn how to learn a task, e.g. online or from little available data.</p> <p>Knowledge and understanding:  At the end of the course students will be familiar with state-of-the-art DNN models for multiple tasks and multi-task objectives, as well as generalization and the effective use of labelled and unlabelled data for learning, self-supervision and meta-learning.</p> <p>Apply knowledge and understanding:  At the end of the course students will have become familiar with the most recent advances in machine learning across a variety of tasks, their adaptation to novel domains and the continual self-learning of algorithms. They will be able to explain the algorithms and choose the most appropriate techniques for a given problem. They will be able to experiment with existing implementations and design and write programs for new solutions for a given task or problem in the two fields.</p> <p>Critical and judgment skills:  Students will be able to analyse a problem or task and identify the most suitable methodologies and techniques to apply in terms of the effective resolution of the problem (accuracy) and its feasibility, including the efficiency, the required amount of data and annotation. Further to class discussions, critical and judgemental skills would be the result of assignments, a course project and a final project report.</p> <p>Communication skills:  Students will acquire the ability to expose their knowledge in a clear and organized way, which will be verified through a final project presentation and its discussion.  Students will be able to express their solutions rigorously and to explain the structure of the code they have written.</p> <p>Learning ability:  The acquired knowledge will enable students to face the study of other problems in machine learning and computer vision. Learning ability would result from the chosen lecture topics, covering most broad areas in advanced machine learning, as well as from the final project, for which students would deep dive into a new topic, beyond the thought material.</p>				
10600503   DATA-DRIVEN MODELING OF COMPLEX SYSTEMS	2°	1°	6	ENG

Insegnamento	Anno	Semestre	CFU	Lingua
<b>Obiettivi formativi</b>				
<p><b>General</b></p> <p>This course aims to exploit advanced techniques from network science and complex systems to understand and eventually predict social-relevant issues (information diffusion, mobility, etc.).</p> <p>The course aims to design efficient strategies to extract knowledge from data through the complex systems approach by stressing the combination of network science and complex systems to build sound mathematical models of complex phenomena.</p> <p>The course will introduce advanced topics of networks science and diffusion models and address the data-driven modeling of complex socio-technical systems (e.g., misinformation diffusion, echo chambers formation, bot detection, mobility patterns, system resilience).</p> <p>The first part of the course will explore the foundational aspects of advanced topics of complex networks (multilayer networks, percolation theory, time-varying graphs). The second part will apply those concepts to actual cases from up-to-date scientific findings ranging from the effect of feed algorithms on social dynamics to patterns of human mobility, passing through information operations, and bot detection.</p> <p>We will use data from real case scenarios (from Facebook, Twitter, Mobility Data, etc.) to analyze phenomena and build and validate models of complex phenomena.</p> <p><b>Specific</b></p> <ul style="list-style-type: none"> <li>• Knowledge and understanding: To know and discuss recent advances in the area of data-driven modeling of complex systems, in particular on algorithms and models to understand and eventually predict social dynamics (e.g., information diffusion, polarization)</li> <li>• Applying knowledge and understanding: to know how to apply criteria and techniques for designing a data analysis framework exploiting the theory of complex systems.</li> <li>• Making judgments: to select the most appropriate strategy to cope with the data-driven modeling of complex phenomena</li> <li>• Communication skills: know how to present projects, including design constraints, solutions, and use possibilities.</li> <li>• Learning skills: ability to develop more advanced studies in data-driven modeling of complex systems.</li> </ul>				
1047214   DATA PRIVACY AND SECURITY	2°	1°	6	ENG

Insegnamento	Anno	Semestre	CFU	Lingua
<b>Obiettivi formativi</b>				
<p><b>General Objectives</b>  Ensuring the privacy of personal data, and securing the computing infrastructures, are key concerns when collecting and analyzing sensitive data sets. Example of these data sets include medical data, personal communication, personal and company-wide financial information. The course is meant to cover an overview of modern techniques aimed at protecting data privacy and security in such applications.</p> <p><b>Specific Objectives</b>  The students will learn the basic cryptographic techniques and their application to obtaining privacy of data in several applications, including cloud computing, statistical databases, distributed computation, and cryptocurrencies.</p> <p><b>Knowledge and Understanding</b>  - ) Modern cryptographic techniques and their limitations.  - ) Techniques for achieving privacy in statistical databases.  - ) Techniques for designing cryptographic currencies and distributed ledgers.  - ) Techniques for secure distributed multiparty computation.</p> <p><b>Applying knowledge and understanding:</b>  - ) How to select the right cryptographic scheme for a particular application.  - ) How to design a differentially private mechanism.  - ) How to program a secure cryptosystem, or a secure smart contract, or a secure cryptographic protocol.</p> <p><b>Autonomy of Judgment</b>  The students will be able to judge the security of the main cryptographic applications.</p> <p><b>Communication Skills</b>  How to describe the security of cryptographic standards, privacy-preserving statistical databases, and blockchains.</p> <p><b>Next Study Abilities</b>  The students interested in research will learn what are the main open challenges in the area, and will obtain the necessary background for a deeper study of the subjects.</p>				
10610252   SIGNAL PROCESSING FOR MACHINE LEARNING	2 <sup>o</sup>	1 <sup>o</sup>	6	ENG

Insegnamento	Anno	Semestre	CFU	Lingua
<b>Obiettivi formativi</b>				
<p><b>Obiettivi</b></p> <p>L'obiettivo del corso è insegnare le metodologie di base dell'elaborazione dei segnali e mostrare la loro applicazione al machine learning e alla data science. I metodi includono: (i) strumenti standard per l'elaborazione di serie temporali e immagini, come l'analisi in frequenza, il filtraggio e il campionamento; (ii) Modelli di dati basati su sparsità e basso rango con applicazioni all'elaborazione di dati con alta dimensionalità (p.es., ricostruzione sparsa, fattorizzazione di matrici, completamento di tensori); (iii) Strumenti di elaborazione di segnali su grafo, adatti ad analizzare ed elaborare dati definiti su domini non metrici (ad es. grafi, ipergrafi, topologie, ecc.) con l'obiettivo di realizzare task di graph machine learning come filtraggio su grafo, clustering spettrale, inferenza della topologia dai dati e reti neurali su grafo. Infine, viene mostrato come formulare e risolvere problemi di machine learning in modo distribuito, adatto per applicazioni di big data, dove l'apprendimento e l'elaborazione dei dati devono essere necessariamente eseguiti su più macchine. Homework ed esercitazioni su dati reali saranno svolti utilizzando Python e/o Matlab.</p> <p><b>Obiettivi specifici:</b></p> <p>6. Conoscenza e comprensione: Apprendere le basi dell'elaborazione dei segnali per il machine learning e applicare questi concetti a problemi di data science.</p> <p>7. Applicazione: Applicare tecniche di elaborazione dei segnali e machine learning su set di dati reali, utilizzando linguaggi di programmazione come Python e Matlab.</p> <p>8. Autonomia di giudizio: Analizzare i vantaggi e i limiti dei diversi strumenti e modelli e determinare la migliore metodologia da utilizzare per un determinato problema.</p> <p>9. Comunicazione: Comunicare in modo efficace nei campi dell'elaborazione dei segnali e del machine learning, considerando la teoria, i vincoli, le soluzioni e le potenziali applicazioni.</p> <p>10. Abilità di apprendimento: Sviluppare studi nel campo dell'elaborazione dei segnali per il machine learning, inclusa la capacità di intraprendere ricerche in questo settore.</p>				
1044406   BIG DATA COMPUTING	2°	1°	6	ENG

Insegnamento	Anno	Semestre	CFU	Lingua
<p><b>Obiettivi formativi</b></p> <p>Obiettivi generali:</p> <p>Conoscenza dei principali scenari applicativi di interesse nell'analisi di collezioni di dati di grandi dimensioni.</p> <p>Conoscenza e comprensione dei principali problemi metodologici e di analisi posti dalla dimensione crescente dei dati.</p> <p>Conoscenza delle principali tecniche di soluzione e dei principali strumenti a disposizione per implementarle.</p> <p>Comprensione degli aspetti teorici e fondazionali delle principali tecniche per l'analisi di collezioni di dati di grandi dimensioni</p> <p>Capacità di tradurre le nozioni acquisite in programmi per la soluzione di problemi specifici.</p> <p>Conoscenza delle principali tecniche di valutazione e loro applicazione a scenari specifici.</p> <p>Obiettivi specifici:</p> <p>Capacità di:</p> <ul style="list-style-type: none"> <li>- individuare le tecniche più adatte a un problema di analisi di dati di grandi dimensioni;</li> <li>- implementare la soluzione proposta, individuando gli strumenti più adatti a raggiungere lo scopo tra quelli disponibili;</li> <li>- progettare e realizzare scenari sperimentali per valutare le soluzioni proposte in condizioni realistiche;</li> </ul> <p>Conoscenza e comprensione:</p> <ul style="list-style-type: none"> <li>- conoscenza dei principali scenari applicativi;</li> <li>- conoscenza delle principali tecniche di analisi;</li> <li>- comprensione dei presupposti teorici e metodologici alla base delle tecniche principali</li> <li>- conoscenza e comprensione delle principali tecniche e indici di valutazione delle prestazioni</li> </ul> <p>Applicare conoscenza e comprensione:</p> <ul style="list-style-type: none"> <li>- essere in grado di tradurre esigenze applicative in problemi concreti di analisi dei dati;</li> <li>- essere in grado di identificare gli aspetti del problema, se presenti, che potrebbero rendere la dimensione (o dimensionalità) dei dati un fattore critico;</li> <li>- essere in grado di individuare le tecniche e gli strumenti più adatti alla soluzione dei problemi concreti di cui sopra;</li> <li>- essere in grado di stimare a priori, almeno qualitativamente, la scalabilità delle soluzioni proposte;</li> </ul> <p>Capacità critiche e di giudizio:</p> <p>Essere in grado di valutare, anche sperimentalmente, l'efficacia, l'efficienza e la scalabilità delle soluzioni proposte</p> <p>Capacità comunicative:</p> <p>Essere in grado di descrivere in modo efficace le specifiche di un problema e di comunicare ad altri le scelte adottate e le motivazioni sottostanti a tali scelte.</p> <p>Capacità di apprendimento:</p> <p>Il corso consentirà lo sviluppo di capacità di approfondimento autonomo su argomenti del corso o ad essi correlati. Metterà lo studente nelle condizioni di individuare e consultare in modo critico manuali avanzati o letteratura scientifica per affrontare scenari nuovi oppure per applicare tecniche alternative a scenari noti.</p>				
1056023   SMART ENVIRONMENTS	2°	2°	6	ENG

**Obiettivi formativi****GENERALI**

L'obiettivo di questo corso è fornire una panoramica del vasto mondo delle tecnologie wireless e cablate che verranno utilizzate per gli ambienti intelligenti. Queste tecnologie saranno in grado di fornire infrastrutture di rete e piattaforme per l'elaborazione delle informazioni digitali utilizzate in ambienti urbani e in ambienti intelligenti.

I recenti progressi in settori quali quelli dell'edge computing, dell'apprendimento automatico, delle reti wireless e rete di sensori consentono varie applicazioni ambientali intelligenti nella vita di tutti i giorni. L'obiettivo principale di questo corso è presentare e discutere i recenti progressi nell'area dell'Internet of Things, in particolare su tecnologie, architetture, algoritmi e protocolli per ambienti intelligenti con enfasi sulle applicazioni reali di ambienti intelligenti. Il corso presenterà gli aspetti di comunicazione e networking, nonché l'elaborazione dei dati da utilizzare per la progettazione dell'applicazione. Il corso proporrà due casi di studio nel campo degli ambienti intelligenti: monitoraggio del traffico veicolare per applicazioni ITS e reti wireless a basso consumo energetico. In entrambi i casi verranno forniti strumenti, modelli e metodologie per la progettazione di applicazioni per ambienti intelligenti.

**SPECIFICI**

- Conoscenza e capacità di comprensione: Conoscere i recenti progressi nell'area dell'Internet delle cose, in particolare su tecnologie, architetture, algoritmi e protocolli per ambienti intelligenti con enfasi sulle applicazioni e sulle piattaforme di elaborazione.
- Capacità di applicare conoscenza e comprensione: saper applicare criteri e tecniche di progettazione di piattaforme intelligenti per l'acquisizione dei dati, per la comunicazione in rete e per le applicazioni in contesti di ambienti intelligenti.
- Autonomia di giudizio: saper analizzare benefici e limiti di progetti ambienti intelligenti.
- Abilità comunicative: saper presentare progetti su ambienti intelligenti e di IoT, compresi vincoli progettuali, soluzioni e possibilità d'impiego.
- Capacità di apprendimento: capacità di sviluppare studi più avanzati nell'ambito delle tecnologie di elaborazione e di rete in ambienti intelligenti.

10616532 |

ECONOMICS AND  
COMPUTATION

2°

2°

6

ENG

**Obiettivi formativi****Obiettivi generali:**

Il corso presenterà un'ampia panoramica di argomenti all'intersezione di informatica, scienza dei dati ed economia, sottolineando l'efficienza, la robustezza e le applicazioni ai mercati online emergenti. Introdurrà i principi della teoria algoritmica dei giochi e della progettazione dei meccanismi economici, della progettazione algoritmica del mercato, nonché dell'apprendimento automatico nei giochi e nei mercati. Dimostrerà applicazioni a casi di studio nella ricerca sul Web e nella pubblicità online, nell'economia delle reti, nei dati, nelle criptovalute e nei mercati dell'intelligenza artificiale.

**Risultati specifici:****Conoscenza e comprensione:**

I principi algoritmici e matematici dell'economia alla base della progettazione e del funzionamento di mercati online efficienti e robusti. L'applicazione di questi principi in esempi concreti di mercati online.

**Applicare conoscenza e comprensione:**

Essere in grado di progettare e analizzare algoritmi per concrete applicazioni dei mercati online rispetto ai requisiti di efficienza e robustezza.

**Capacità critiche e di giudizio:**

Essere in grado di valutare la qualità di un algoritmo per applicazioni nel mercato online, discriminando gli aspetti di modellizzazione da quelli legati all'implementazione algoritmica e di sistema.

**Capacità comunicative:**

Capacità di comunicare e condividere le scelte di modellazione e i requisiti di sistema, nonché i risultati dell'analisi dell'efficienza degli algoritmi del mercato online.

**Capacità di apprendimento:**

Il corso stimola gli studenti ad acquisire capacità di apprendimento al crocevia tra informatica, economia e applicazioni del mercato digitale, compresi i diversi linguaggi utilizzati in questi campi.

Insegnamento	Anno	Semestre	CFU	Lingua
10616533   GRAPH MINING AND APPLICATIONS	2°	2°	6	ENG

### Obiettivi formativi

Risultati di apprendimento attesi:

Graphs have applications in multiple areas, including social networks, bioinformatics, network medicine, computational chemistry, and they can be used to provide tools in these areas.

The course will present models and algorithms for the analysis of graphs as with applications on various areas. The goal at the end of the course, is for student to know algorithms and frameworks that can allow them to analyze large graph data.

Informazioni sui prerequisiti culturali e curriculari necessari

- Knowledge of basic algorithms
- Programming
- Linear algebra
- Probability
- Neural networks

Programma in italiano

- Theoretical algorithms for graph modeling and analysis:
  - ? Real graph properties and models (Gnp, preferential attachment, Kleinberg's reachability)
  - ? Models for propagation (linear threshold, cascade) and for opinion formation
  - ? Homophily and influence and algorithms for identifying and distinguishing
  - ? Influence maximization
  - ? Algorithms for graph alignment
  - ? Dense subgraphs, community detection, graph minors
  - ? Graph summarization and sampling
- Machine-learning approaches:
  - ? Label propagation
  - ? Graph transformers
  - ? Knowledge-graph emdeddings
  - ? Models for analysis of temporal graphs
  - ? Explainability
- Architectures for handling large graph data:
  - ? Spark GraphsX
  - ? AWS Neptune
  - ? AWS GraphStorm
  - ? Neo4J

Modalità di valutazione delle conoscenze

- Prova scritta
- Prova orale
- Valutazione progetto

Modalità di valutazione in italiano

Homeworks and/or project and oral exam or written exam

Esempi di domande e/o esercizi frequenti

Find the most influential nodes in a network.

Testi adottati

Material will be distributed online

Modalità di svolgimento

Didattica frontale/tradizionale

Modalità di svolgimento in italiano

The course is based on in-class theoretical lectures and sometimes in-class labs.

Modalità di frequenza

Classes are in person.

Programmazione:

<http://aris.me/index.php/teaching>

Insegnamento	Anno	Semestre	CFU	Lingua
1047209   QUANTITATIVE MODELS FOR ECONOMIC ANALYSIS AND MANAGEMENT	1°	2°	6	ENG

### Obiettivi formativi

General Objectives of the course

The general objectives of the course are:

- Present a general framework for the development of quantitative models for economic analysis and management;
- Provide the basic concepts and a guide to analyse the specialised literature;
- Propose a unified framework on the main methodologies available to compare the productivity and efficiency of Decision Making Units (DMUs);
- Introduce to the relevant roles played by the data for the development of effective quantitative models of socio-economic systems;
- Make an introduction to the main softwares available to implement the quantitative models presented during the course;
- Provide laboratory sessions to implement the quantitative models presented during the course in practice;
- Present several applications in the field of economics and management, including public sector services as potential group project works, to be developed by the students according to their personal interest and background;
- Interact with students through seminars, assisted laboratory, oral presentations and the realization of a project work on real data.

Specific objectives of the course

- KNOWLEDGE AND UNDERSTANDING: DEMONSTRATE THE KNOWLEDGE OF THE BASIC METHODS FOR THE DEVELOPMENT OF QUANTITATIVE MODELS FOR ECONOMIC ANALYSIS AND MANAGEMENT ;
- ABILITY TO APPLY KNOWLEDGE AND UNDERSTANDING: TO BE ABLE TO DEVELOP QUANTITATIVE ECONOMIC MODELS ON THE BASE OF THE KNOWLEDGE AND TECHNIQUES LEARNED DURING THE COURSE;
- JUDGMENT AUTONOMY: TO BE ABLE TO DEVELOP A QUANTITATIVE ECONOMIC MODEL WITH CRITICAL SPIRIT, CHOOSING THE APPROPRIATE METHOD AND CORRECTLY IMPLEMENTING IT.
- COMMUNICATION SKILLS: BEING ABLE TO COMMUNICATE THE RESULTS OF THE ANALYSIS AND ITS INFORMATION TO DIFFERENT TYPES OF INTERLOCUTORS;
- LEARNING SKILLS: TO DEVELOP THE NECESSARY SKILLS TO APPLY AND DEVELOP AUTONOMOUSLY THE METHODS AND MODELS LEARNED DURING THE COURSE.

10600197   Data Driven Economics	1°	2°	6	ENG
-------------------------------------	----	----	---	-----

Insegnamento	Anno	Semestre	CFU	Lingua
<b>Obiettivi formativi</b>				
1) Knowledge and understanding During the lectures of Data-driven Economics, students acquire the basic theoretical elements of econometric analysis. Theoretical lectures are aimed at guiding students in the acquisition of the basics of simple and multiple regression models, starting from the relative assumptions, and then proceeding with the estimation and inference procedures. The course contents cover both the estimation of linear and non-linear models and the analysis of both cross-sectional and longitudinal data.				
2) Applying knowledge and understanding The students of the Data-driven Economics course are able to apply the notions acquired during the theoretical lectures to a wide range of problems of an empirical nature. They acquire the ability to build econometric models aimed at giving empirical content to economic relations and are also able to establish a causal link between two or more variables in the economic field.				
3) Making judgements Students are encouraged to critically discuss empirical studies published in the economic/managerial field in the classroom. The Data-driven Economics course also includes a laboratory in which students apply the acquired knowledge of econometrics to the estimation of empirical models carried out using data made available by the teacher.				
4) Communication skills At the end of the course, students are able to illustrate and explain the strengths and weaknesses of a wide range of empirical methodologies to a variety of heterogeneous interlocutors in terms of training and professional role. The acquisition of these skills is verified and evaluated not only during the final exam, by means of a written test and a possible oral test, but also during flipped class sessions in which, individually or in groups, students are called to present empirical studies published in the economic/managerial field.				
5) Learning skills Students acquire the ability to independently conduct empirical analyses by building econometric models to be estimated using data with diversified structures. The tools provided by the course allow for the analysis of systems in which a large number of factors simultaneously contribute to explaining their states and impact assessments that take into account the uncertainty and risk inherent in the application of policies. The acquisition of these skills is verified and evaluated during the final exam, by means a written test and a possible oral test, in which the student can be called to discuss empirical problems on the basis of the topics covered and the reference material distributed during the course.				

Lo studente deve acquisire 12 CFU fra i seguenti esami

Insegnamento	Anno	Semestre	CFU	Lingua
1056085   BIG DATA FOR OFFICIAL STATISTICS	2 <sup>o</sup>	1 <sup>o</sup>	6	ENG
<b>Obiettivi formativi</b>				
What subset of Big Data can be used in the ambit of Official Statistics and what domains of Official Statistics can be enriched through the availability of new data sources. How new data sources can be used in Official Statistics, by taking into account challenges, needs and risks in this exercise.				
Definition of the role of Big Data in the context of Official Statistics. How to frame the measurement of social, demographic and economic phenomena through Big Data by considering challenges, needs and risks.				
10589627   NEURAL NETWORKS FOR DATA SCIENCE APPLICATIONS	2 <sup>o</sup>	1 <sup>o</sup>	6	ENG

**Insegnamento****Anno****Semestre****CFU****Lingua****Obiettivi formativi**

General objectives: The course provides an overview on the use of deep neural networks in the context of data science and data science applications. The course is split into a methodological part (introducing basic concepts and tools for building neural networks), and a practical part with several hands-on coding sessions, followed by one homework, one final project, and an oral examination.

Specific objectives: The first part of the course will (briefly) reintroduce the mathematical skills required for the course, including linear algebra and numerical optimization. Then, we will survey basic neural network components ranging from linear models to fully-connected ones layers. We will then move to a selection of advanced models (convolutive networks, transformers, graph neural networks, autoregressive models), and a series of selected advanced topics (fairness, robustness, deployment of the models).

Knowledge and understanding: At the end of the course, the students will have a broad knowledge of state-of-the-art tools and techniques for implementing deep neural networks in several fields, as long as practical hands-on ability to translate conceptual designs into practical coding.

Critical and judgment skills: The students will learn to tackle a complex data science project, decomposing it into blocks that are solvable through one or more neural network models.

Communication skills: The students will learn to effectively communicate their knowledge along three major axes, (i) via suitably describing their final projects with a final report, (ii) orally for the final exam, and (iii) through careful code documentation and restructuring.

Learning ability: The students will be able to autonomously read and reimplement state-of-the-art papers and models going beyond the basic topics of the course, thanks to a selection of papers and tools that will be discussed during the lectures.

10593052 |  
BIOINFORMATICS  
AND NETWORK  
MEDICINE

2°

1°

6

ENG

**Obiettivi formativi**

General objectives. The general objectives of the course are: i) to provide students with a hands-on experience with basic biological concepts and common bioinformatics tools and databases; ii) to introduce students to the on-the-field application of networks in biology and medicine.

Specific objectives. Students are expected to acquire basic biology knowledge and skills, to understand the role of networks in the study of physiological mechanisms and diseases; to understand how to use network medicine algorithms and procedures.

Knowledge and understanding. The course will include theory and hands-on projects. Students will be trained in the basic theory and application of programs used for database searching, biological network inference and analysis.

Apply knowledge and understanding. At the end of the course students will have become familiar with basic biological concepts and bioinformatics databases and tools. Furthermore, on successful completion of this course, students will understand the use of networks as a paradigm for disease expression and course.

Critical and judgment skills. At the end of the course, students will be able to critically analyse the results of their analysis.

Communication skills. The students will be required to produce reports describing the hands-on projects with specific sections for the description of the obtained results and their discussion.

Learning ability. The projects will be developed in small groups encouraging team building. All the acquired abilities will be checked in a final oral exam during which a good division of teamwork will be rewarded.

10593053 | DIGITAL  
EPIDEMIOLOGY  
AND PRECISION  
MEDICINE

2°

1°

6

ENG

**Insegnamento****Anno****Semestre****CFU****Lingua****Obiettivi formativi**

General objectives. Digital data sources and digital traces of human behaviour have the potential to provide local and timely information about disease and health dynamics at the population level. The general aim of the course is to introduce students to the analysis of epidemiological and omics data and to the use of computational approaches for medical/clinical purposes.

Specific objectives. The course consists of two modules. The first module will deal with the opportunities and challenges of mining digital data sources for epidemiological and public health signals and will provide an overview of the state of the art of this emerging field. The second module will focus on "precision medicine", an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person. With the second module, the students are expected to acquire basic biology knowledge and skills and to become familiar with the analysis and integration of omics data.

Knowledge and understanding. The course will include theory and hands-on lectures. Students will be trained in the basic theory for the identification of gene interactions and in the use of network science.

Apply knowledge and understanding. At the end of the course students will have become familiar with basic biological concepts, with the analysis of omics and epidemiological data and with the use of networks for the investigation of infectious disease dynamics and disease etiology, diagnosis, and treatment.

Critical and judgment skills. At the end of the course, students will be able to critically analyse the results of their analysis.

Communication skills. The students will be required to produce reports describing hands-on projects with specific sections for the description of the obtained results and their discussion.

Learning ability. The projects will be developed in small groups encouraging team building.

1047212 |

Economics of

2<sup>o</sup>2<sup>o</sup>

6

ENG

Network Industries

**Obiettivi formativi**

Knowledge and understanding

The aim of the course is to introduce students to the new information economy and the economics of network industries.

Students are expected to gain insight into how the specific features of technology and demand affect market structure, firms' strategies and business models, as well as public policy in network industries.

Applying knowledge and understanding

By the end of the course, students should be able to use methods and models of microeconomics and industrial organization to understand and analyze the competitive dynamics in the new information economy, and specifically in network industries.

Making judgements

Lectures, practical exercises and problem-solving sessions will provide students with the ability to assess the main strengths and weaknesses of theoretical models when used to explain empirical evidence and case studies in the new information economy.

Communication

By the end of the course, students are able to point out the main features of the new information economy and network industries, and to discuss relevant information, ideas, problems and solutions both with a specialized and a non-specialized audience. These capabilities are tested and evaluated in the final written exam and possibly in the oral exam as well as in the project work.

Lifelong learning skills

Students are expected to develop those learning skills necessary to undertake additional studies on relevant topics in the field of the new information economy with a high degree of autonomy. During the course, students are encouraged to investigate further any topics of major interest, by consulting supplementary academic publications, specialized books, and internet sites. These capabilities are tested and evaluated in the final written exam and possibly in the oral exam as well as in the project work, where students may have to discuss and solve some new problems based on the topics and material covered in class.

1047222 |

EFFICIENCY AND  
PRODUCTIVITY  
ANALYSIS2<sup>o</sup>2<sup>o</sup>

6

ENG

Insegnamento	Anno	Semestre	CFU	Lingua
<b>Obiettivi formativi</b>				
<p>This course has the target of providing the students with the modern techniques of measuring quantitatively advanced topics in economic statistics. In particular our focus will be on three main interrelated directions: 1) the analysis of production and efficiency, specifically in the private but also in the public sectors, 2) economic dynamics of sectorial systems founded on micro data, 3) growth, ICT and technology in the modern economy.</p> <p>This course uses statistical methods, both stochastic and deterministic, to analyze topics such as productivity, efficiency and growth at micro, sectorial, and for coherence at macro level. We first take into exam data from firms that will be useful for the mentioned three-levels study, then, as regards the efficiency analysis of productive units, such data will be employed in order to evaluate mergers and acquisitions of plants and firms and management of productive factors. Efficiency will be evaluated from the sides of costs, profits and revenues. As for the sectorial analysis, static and dynamic models will be considered to allow for forecasts and simulations in each sector for variables like production, labour, capital, raw materials, prices and capital gains. As a consequence, an aggregate analysis on the production, growth and prices will follow. We also deal with ICT and technical progress in the production process considering how and if the associated externalities are effective. We will use the following techniques for data analysis: accounting rules for the database, panel data econometrics, time series analysis for systems of equations, methods for differential equation systems. Topics on private and also public sectors will contribute to explain the relationship between economic structure and the actual crisis. Specifically, lectures also include the examinations of cases study concerning the efficiency and productivity analysis on the recent patterns of the banking sector in the international context.</p>				
10589730   GEOMATICS AND GEOINFORMATION	2°	2°	6	ENG

#### Obiettivi formativi

The course finds its motivation in the great availability and relevance of geospatial data (in particular big data), and it aims to provide the fundamentals on the main methodologies and techniques currently available for their acquisition, verification, analysis, storage and sharing.

In fact, the vast majority (a percentage close to 80%) of the currently available data has a geographical connotation, is intrinsically linked to a position; they are therefore named geospatial data. Furthermore, the ever-increasing availability of sensors capable of acquiring geospatial data, allowing the acquisition of larger and larger amounts of data, raises several important issues related to the correct, efficient and effective use of these geospatial big data.

Special attention is given to data coming from Global Navigation Satellite Systems (GNSS), Photogrammetry and Remote Sensing, Volunteered Geographic Information (VGI) and crowdsourcing, both regarding their analysis and management with freely available software and cloud-based platforms for planetary-scale environmental data analysis (Google Earth Engine).

#### Knowledge and understanding

Students who have passed the exam will know the fundamentals on the main methodologies and techniques currently available for geospatial data acquisition, verification, analysis, storage and sharing, with focus on Global Navigation Satellite Systems (GNSS), Photogrammetry and Remote Sensing, and cloud-based platforms for planetary-scale environmental data analysis (Google Earth Engine), being also aware of the relevant resources represented by Volunteered Geographic Information (VGI) and crowdsourcing

#### Applying knowledge and understanding

Students who have passed the exam will be able to plan and manage the acquisition, verification, analysis, storage and sharing of geospatial data necessary to solve interdisciplinary problems, using Global Navigation Satellite Systems (GNSS), Photogrammetry and Remote Sensing, and cloud-based platforms for planetary-scale environmental data analysis (Google Earth Engine), being also aware of the relevant additional contributions which can be supplied by Volunteered Geographic Information (VGI) and crowdsourcing

#### Making judgment

Students will acquire autonomy of judgment thanks to the skills developed during the execution of the numerical and practical exercises that will be proposed on three main topics of the course (Global Navigation Satellite Systems, Photogrammetry and Remote Sensing, Google Earth Engine)

#### Learning skills

The acquisition of basic methodological skills on the topics covered, together with state-of-the-art operational skills, favors the development of autonomous learning skills by the student, allowing continuous, autonomous and thorough updating.

Insegnamento	Anno	Semestre	CFU	Lingua
1047218   EARTH OBSERVATION DATA ANALYSIS	2°	2°	6	ENG

#### Obiettivi formativi

The module aims at providing a general background on the remote sensing systems for Earth Observation from spaceborne platforms and on data processing techniques. It describes, using a system approach, the characteristics of the system to be specified to fulfil the final user requirements in different domains of application. Remote sensing basics and simple wave-interaction models useful for data interpretation are reviewed together with technical principles of the main remote sensors. The course also provides an overview of the most important applications and biogeophysical parameters (of the atmosphere, the ocean and the land) which can be retrieved. The most important techniques for data processing and product generation, also by proposing practical exercises using the computer, are analysed together with an overview of the main Earth Observation satellite missions and the products they provide to the final user.

1047215   INTELLECTUAL PROPERTY COMPETITION AND DATA PROTECTION LAW	2°	2°	6	ENG
--	----	----	---	-----

#### Obiettivi formativi

The aim of the course is to provide students with an overview of the functioning of intellectual property, competition and data protection law from both an economic and legal perspective. By the end of the course students are expected to have acquired a general understanding of the main policy issues involved, and should be able to identify and apply the relevant legal rules, both substantial and procedural, in situations that can be considered routinary to professionals and businesses operating in the data science industry.

## Obiettivi formativi

Il corso di studio in Data Science si caratterizza per un'offerta didattica interdisciplinare che raccoglie contributi dell'ingegneria, dell'informatica, della statistica, delle scienze economiche e organizzative, insieme a conoscenze specifiche dei principali domini applicativi di Data Science. In particolare, la laurea magistrale in Data Science offre le conoscenze professionali adeguate per lo sviluppo delle tecnologie di raccolta, gestione, elaborazione e analisi dei big data, e la conseguente traduzione in informazioni fondamentali per il processo conoscitivo e decisionale all'interno dei settori innovativi di business e sociali. La laurea magistrale in Data Science si pone

l'obiettivo della formazione di nuove figure professionali che possano contribuire ad aumentare l'efficienza ed affidabilità delle istituzioni pubbliche, delle aziende private e delle amministrazioni locali, con particolare riferimento agli open data, al loro utilizzo per lo sviluppo di servizi più efficienti per le aziende e i cittadini e per l'ottimizzazione della gestione delle risorse nei contesti urbani. La laurea magistrale in Data Science si pone come obiettivo anche la formazione di professionisti in grado di operare all'interno delle agenzie pubbliche e private al fine di integrare i big data all'interno dei processi di analisi economico e sociale. Le caratteristiche di interdisciplinarietà della laurea magistrale in Data Science e la sua rigorosa impostazione metodologica la rendono adatta ad essere fruita da studenti che abbiano conseguito la laurea di primo livello in tutti i settori dell'Ingegneria dell'Informazione, dell'Informatica e della Statistica, nonché nelle facoltà di Economia, Matematica e Fisica. Oltre alle conoscenze specifiche del settore, costituiscono parti fondamentali dell'offerta formativa gli aspetti teorico-scientifici necessari a descrivere e a interpretare i problemi del contesto applicativo in cui si pone il problema di sviluppare metodologie innovative di Data Science, lo sviluppo di capacità di ideazione, pianificazione, progettazione e gestione di complessi sistemi di gestione e analisi di grandi moli di dati, lo sviluppo di capacità di sperimentazione, la conoscenza e l'uso fluente della lingua inglese. Costituisce un elemento di completamento essenziale della formazione la prova finale o tesi di laurea magistrale, che permette al laureando di applicare la pluralità di nozioni e metodologie acquisite in un campo di applicazione industriale, scientifico o di analisi economico-sociale, e che dimostra la padronanza degli argomenti, la capacità di operare in modo autonomo e un buon livello di comunicazione. Il percorso formativo è orientato alla fruibilità della laurea magistrale in ambito internazionale, fruibilità garantita anche dall'erogazione in lingua inglese. Il percorso formativo è inoltre orientato a mantenere una stretta connessione con il tessuto lavorativo. Il laureato magistrale in Data Science avrà anche un livello di preparazione adeguato per una sua collocazione in contesti di ricerca sia di base che applicata, sia presso università e centri di ricerca che presso settori aziendali di ricerca e sviluppo, sia in ambito nazionale e internazionale. Il regolamento didattico del corso di studio definirà, nel rispetto dei limiti normativi, la quota dell'impegno orario complessivo a disposizione dello studente per lo studio personale e per altre attività formative di tipo individuale. La proposta formativa prevede un primo insieme di insegnamenti su settori scientifico disciplinari caratterizzanti obbligatori per tutti gli studenti nel primo anno di corso. Di questi insegnamenti obbligatori una parte comprende corsi nei SSD INF/01 Informatica, ING-INF/05 Sistemi di elaborazione delle informazioni e ING-INF/03 Telecomunicazioni. Questo insieme di insegnamenti obbligatori mira a garantire l'omogeneizzazione delle competenze e conoscenze in ambito informatico e dell'informazione considerata l'eterogeneità dei percorsi formativi triennali di provenienza degli studenti, italiani e stranieri. Attraverso questi insegnamenti si forniscono le conoscenze informatiche e dell'ingegneria dell'informazione di base nell'ambito della programmazione avanzata, degli algoritmi per acquisire conoscenza sui dati, delle librerie per la visualizzazione e l'analisi dei dati, del machine learning, e delle infrastrutture di rete per l'immagazzinamento, l'elaborazione, e la comunicazione delle grandi moli di dati. L'altra quota parte di questi insegnamenti obbligatori è invece orientata a irrobustire le fondamenta nell'area dei modelli matematico statistici utili per l'analisi di grandi moli di dati e comprende corsi di SSD matematici (MAT/06-09) e statistici (SECS-S/01). Tutti questi insegnamenti obbligatori includono anche attività di laboratorio e attività progettuali autonome. Al primo anno si aggiunge un insegnamento a scelta in un gruppo caratterizzante nell'ambito delle discipline umane, sociali, giuridiche ed economiche e un altro insegnamento a scelta nel gruppo caratterizzante nell'ambito della formazione matematico-statistica e della probabilità. Sempre al primo anno lo studente dovrà scegliere almeno un insegnamento in un gruppo caratterizzante relativo alla formazione informatica e dell'informazione. Nel secondo anno gli studenti includeranno altri insegnamenti nell'ambito della formazione informatica e dell'informazione. Il percorso formativo prevede complessivamente un'ampia scelta di insegnamenti all'interno delle attività caratterizzanti in ambito informatico e dell'informazione permettendo allo studente di acquisire competenze specifiche che possono spaziare dal machine learning e data mining avanzati, all'applicazione di tecniche computazionali per l'analisi e la sintesi del linguaggio naturale, alle reti sociali, all'elaborazione dei segnali su grafi e agli aspetti fondamentali sull'analisi di sistemi complessi. In quest'area gli studenti si potranno altresì specializzare in argomenti relativi alla privacy e alla sicurezza dei dati, alla raccolta ed elaborazione dati in ambienti smart e alla memorizzazione e gestione di grandi moli di dati in database e infrastrutture cloud. Il percorso formativo al secondo anno prevede ulteriori attività formative che includono dei training camp tematici organizzati in collaborazione con aziende del settore o tirocini formativi e insegnamenti scelti nei settori scientifici disciplinari affini. Sono poi previste le attività formative autonomamente scelte dello studente e la prova finale e la stesura della tesi di laurea magistrale.

## **Profilo professionale**

### **Profilo**

Big Data Infrastructure Professional

## **Funzioni**

Il Big Data Infrastructure Professional deve essere in grado di installare e gestire le più importanti infrastrutture software, hardware e di rete per i big data. Il suo compito è quello di garantire il funzionamento dell'infrastruttura hardware e software e di consigliare le figure professionali manageriali riguardo la scelta, e l'utilizzo delle infrastrutture più adatte alle situazioni contingenti.

## **Competenze**

Il Big Data Infrastructure Professional sarà in grado di: - Progettare le architetture hardware, software e di rete per i big data; - Gestire l'operatività delle infrastrutture per i big data e comprendere problematiche relative al loro funzionamento e alla loro prestazioni; - Identificare strumenti per la valutazione di aspetti legati alla sicurezza informatica delle infrastrutture e dei dati in esse trasferiti e processati;

## **Sbocchi lavorativi**

Grandi aziende, per esempio operatori nell'ambito di servizi Cloud; Pubblica amministrazione, per esempio per la dotazione di servizi di storage e di cloud computing.

# Frequentare

## Laurearsi

La prova finale potrà essere inerente a un'attività progettuale, di ricerca, metodologica o di tirocinio, presso una struttura industriale, istituzioni pubbliche di ricerca o presso i laboratori stessi dell'università. L'esame finale di laurea consiste nella presentazione e discussione di un progetto con caratteri di originalità e di una relazione supervisionata da un docente di riferimento. Il lavoro svolto dovrà dimostrare che lo studente ha raggiunto una padronanza delle metodologie di Data Science e/o della loro applicazione in un settore specifico a un livello di competenza in linea con le esigenze imposte dai processi di innovazione tecnologica. La prova finale sarà impostata in maniera tale da costituire una credenziale importante per l'inserimento del laureato nel tessuto lavorativo. La prova finale si svolgerà mediante una discussione da parte del candidato dinanzi a una Commissione, la cui composizione è stabilita dagli appositi regolamenti di Ateneo. Le Commissioni giudicatrici per la prova finale esprimono la loro votazione in centodecimi e possono, all'unanimità, aggiungere la lode al massimo dei voti.

# Organizzazione

## Presidente del Corso di studio - Presidente del Consiglio di area didattica

Francesca Cuomo

## Tutor del corso

## Manager didattico

Antonella Palombo

## Rappresentanti degli studenti

Susanna Bravi  
Stefano Rinaldi

## Docenti di riferimento

ARISTIDIS ANAGNOSTOPOULOS  
FABIO GALASSO  
PIERPAOLO BRUTTI  
CINZIA DARAIO  
LUCA BECCHETTI  
VERONICA PICCIALLI  
FILOMENA MAGGINO  
STEFANIA COLONNESE

## Regolamento del corso

Obiettivi formativi specifici \*\*\*\*\* Il corso di studio si caratterizza per un'offerta didattica interdisciplinare che raccoglie contributi dell'ingegneria, dell'informatica, della statistica, delle scienze economiche e organizzative, insieme a conoscenze specifiche dei principali domini applicativi di Data Science. In particolare, la laurea magistrale in Data Science offre le conoscenze professionali adeguate per lo sviluppo delle tecnologie di raccolta, gestione, elaborazione e analisi dei big data, e la conseguente traduzione in informazioni fondamentali per il processo conoscitivo e decisionale all'interno dei settori innovativi di business e sociali. Un percorso formativo in Data Science risponde alle notevoli sfide scientifiche e tecnologiche legate all'emergere di piattaforme globali di memorizzazione ed elaborazione dei dati: i dati personali, governativi e commerciali e le relative applicazioni abbandonano i sistemi proprietari per approdare ai sistemi di cloud computing e cloud storage con i relativi problemi di affidabilità, privacy e sicurezza. La laurea magistrale in Data Science si pone l'obiettivo della formazione di nuove figure professionali che possano contribuire ad aumentare l'efficienza ed affidabilità delle istituzioni pubbliche, delle aziende private e delle amministrazioni locali, con particolare riferimento agli open data, al loro utilizzo per lo sviluppo di servizi più efficienti per le aziende e i cittadini e per l'ottimizzazione della gestione delle risorse nei contesti urbani. La laurea magistrale in Data Science si pone come obiettivo anche la formazione di professionisti in grado di operare all'interno delle agenzie pubbliche e private al fine di integrare i big data all'interno dei processi di analisi economico e sociale. Le caratteristiche di interdisciplinarietà della laurea magistrale in Data Science e la sua rigorosa impostazione metodologica la rendono adatta ad essere fruita da studenti che abbiano conseguito la laurea di primo livello in tutti i settori dell'Ingegneria dell'Informazione, dell'Informatica e della Statistica, nonché nelle facoltà di Economia, Matematica e Fisica. Oltre alle conoscenze specifiche del settore, costituiscono parti fondamentali dell'offerta formativa gli aspetti teorico-scientifici necessari a descrivere e a interpretare i problemi del contesto applicativo in cui si pone il problema di sviluppare metodologie innovative di Data Science, lo sviluppo di capacità di ideazione, pianificazione, progettazione e gestione di complessi sistemi per il trattamento e l'analisi di grandi moli di dati, lo sviluppo di capacità di sperimentazione, la conoscenza e l'uso fluente della lingua inglese. Costituisce un elemento di completamento essenziale della formazione la prova finale o tesi di laurea magistrale, che permette al laureando di applicare la pluralità di nozioni e metodologie acquisite in un campo di applicazione industriale,

scientifico o di analisi economico-sociale. Con la stesura della tesi lo studente dimostra la padronanza degli argomenti, la capacità di operare in modo autonomo e un buon livello di comunicazione. Il percorso formativo è orientato alla fruibilità della laurea magistrale in ambito internazionale, fruibilità garantita dalla quantità e dalla qualità delle relazioni internazionali di ricerca facenti capo ai docenti, nonché dall'erogazione in lingua inglese. Il percorso formativo è inoltre orientato a mantenere una stretta connessione con il tessuto lavorativo, connessione garantita da gran numero e prestigio dei progetti di ricerca applicata di cooperazione tra università e aziende nazionali, internazionali e pubblica amministrazione in cui i docenti sono coinvolti. Il laureato magistrale in Data Science avrà anche livello di preparazione adeguato per una sua collocazione in contesti di ricerca sia di base che applicata, sia presso università e centri di ricerca che presso settori aziendali di ricerca e sviluppo, sia in ambito nazionale e internazionale. L'offerta didattica si avvarrà di tutte le competenze multidisciplinari offerte da tutti i 4 Dipartimenti della Facoltà di Ingegneria dell'Informazione, Informatica e Statistica (I3S), il Dipartimento di Scienze Statistiche, il Dipartimento di Ingegneria Informatica, Automatica e Gestionale "Antonio Ruberti", il Dipartimento di Informatica e il Dipartimento di Ingegneria dell'Informazione, Elettronica e Telecomunicazioni.

Requisiti di ammissione e crediti riconoscibili \*\*\*\*\* - Requisiti di ammissione Per essere ammessi al corso di laurea magistrale occorre essere in possesso dei necessari requisiti curriculari (RC) e di un'adeguata preparazione personale (APP). Requisiti curriculari I requisiti curriculari (RC) sono i seguenti: (RC-a) Il possesso della laurea o del diploma universitario di durata triennale, ovvero di altro titolo di studio conseguito all'estero, ritenuto idoneo (RC-b) L'aver acquisito almeno 90 crediti formativi universitari nell'insieme dei seguenti settori: - Scienze matematiche e informatiche: MAT/\*, INF/01 - Scienze fisiche: FIS/\* - Scienze economiche e statistiche: SECS-P/\*, SECS-S/\* - Ingegneria industriale e dell'informazione: ING-IND/\*, ING-INF/\* - Scienze biologiche: BIO/\* - Scienze giuridiche: IUS/\* - Scienze della terra: GEO/\* - Ingegneria civile e ambientale: ICAR/\* (RC-c) Conoscenza della lingua inglese a livello B2 o superiore. Tali requisiti curriculari intendono consentire l'accesso al corso di studio da parte di tutti gli studenti che abbiano conseguito la laurea di primo livello nelle classi di laurea L-8 (Ingegneria dell'informazione), L-31 (Scienze e tecnologie informatiche) e L-41 (Statistica), nonché nelle classi di laurea L-18 (Scienze dell'economia e della gestione aziendale), L-30 (Scienze e tecnologie fisiche), L-33 (Scienze economiche) e L-35 (Scienze matematiche) e nelle corrispondenti classi di cui al D.M. 509/1999. Modalità di ammissione: Per essere ammessi al corso di laurea magistrale occorre essere in possesso dei necessari requisiti curriculari (RC) e di un'adeguata preparazione personale (APP). La verifica dei requisiti per l'ammissione e, in particolare, il possesso della preparazione personale sarà verificato da una apposita Commissione nominata dal Consiglio di Corso di Studio. Verifica dei requisiti curriculari

\*\*\*\*\* Lo studente deve soddisfare contemporaneamente (RC-a), (RC-b) e (RC-c). La verifica del requisito (RC-c) viene soddisfatta da tutti gli studenti in possesso di certificazione della lingua inglese a livello B2 ovvero che certifichino di avere acquisto crediti di lingua inglese (anche idoneità) a livello B2 nella loro carriera pregressa. In assenza di idonea certificazione ovvero dei crediti suddetti gli studenti dovranno superare un colloquio di verifica della conoscenza della lingua inglese. Verifica dell'adeguata preparazione personale

\*\*\*\*\* L'adeguata preparazione personale (APP) considera due aspetti: (APP-a) Risultati e pertinenza della pregressa carriera; (APP-b) Conoscenze riguardo alla Matematica, Probabilità e Informatica. Per la verifica dei requisiti relativi ai risultati e alla pertinenza (APP-a) saranno oggetto di valutazione: la votazione finale ottenuta nella laurea di primo livello e la relativa media dei voti (ovvero esclusivamente la media dei voti nel caso il candidato non fosse ancora laureato) facendo particolare attenzione ai voti conseguiti nell'area Matematica, Probabilità ed Informatica. la pertinenza del curriculum di studio della laurea di primo livello. La verifica dell'adeguata preparazione riguardo alla Matematica, Probabilità e Informatica (APP-b) considererà la conoscenza acquisita sui seguenti argomenti: (APP-b1). Matematica: Calcolo differenziale ed integrale per funzioni di una o più variabili reali; nozioni base di algebra lineare e geometria analitica nel piano e nello spazio. (APP-b2). Probabilità: Variabili aleatorie, distribuzioni e valori attesi; principali modelli di variabili aleatorie; convergenza per successioni di variabili aleatorie. (APP-b3). Informatica: Principi di programmazione, rudimenti di object-oriented design; almeno un linguaggio di programmazione tra C, C++, C#, Java, Python. La Commissione considererà automaticamente soddisfatta la verifica del requisito APP-b per gli studenti che abbiano acquisito almeno: 12 crediti complessivi nei settori MAT/03 (Geometria) e/o MAT/05 (Analisi Matematica), 6 crediti nel settore MAT/06 (Probabilità), 6 crediti complessivi nei settori INF/01 (Informatica) e/o ING-INF/05 (Sistemi di elaborazione delle informazioni). In caso contrario gli studenti dovranno sostenere un test e, se supereranno il test, dovranno superare un colloquio finale di accertamento delle conoscenze (APP-b1, APP-b2, APP-b3) per completare la verifica del possesso dell'adeguata preparazione personale (APP-b).

Descrizione del percorso formativo \*\*\*\*\* La proposta formativa prevede un primo insieme di 39 CFU su settori scientifico disciplinari caratterizzanti obbligatori per tutti gli studenti nel primo anno di corso. Questi CFU sono mirati a fornire le conoscenze statistiche, ingegneristiche e informatiche di base per la raccolta, l'elaborazione, l'organizzazione delle grandi moli di dati e lo sviluppo dei modelli matematico statistici utili per la loro analisi. Tutti questi insegnamenti obbligatori includono anche attività di laboratorio e attività progettuali autonome. Al primo anno si aggiungono 6 CFU a scelta in un gruppo caratterizzante

nell'ambito delle discipline umane, sociali, giuridiche ed economiche e altri 6 CFU a scelta nel gruppo caratterizzante nell'ambito della formazione matematico-statistica. Sempre al primo anno lo studente dovrà scegliere almeno 6 CFU in un gruppo caratterizzante relativo alla formazione informatica e dell'informazione. Nel secondo anno gli studenti includeranno altri CFU nell'ambito della formazione informatica e dell'informazione per raggiungere 18 CFU complessivi in quest'area. Il percorso formativo al secondo anno prevede quindi 3 CFU di ulteriori attività formative (AAF) e 12 CFU scelti nei settori scientifici disciplinari affini. Infine i 12 CFU a scelta libera dello studente, di cui al massimo 6 nel primo anno, completano il percorso. I 3 CFU per Altre Attività Formative (AAF) includono sia tirocini formativi e di orientamento, sia attività di training camp organizzate con la partecipazione di aziende del settore sia attività utili per l'inserimento nel mondo del lavoro. Il percorso formativo si completa con la prova finale (tesi di laurea magistrale) di 24 CFU. Non sono previsti obblighi di frequenza se non per attività di laboratorio e attività pratiche. Tutti gli insegnamenti sono erogati in lingua inglese. Sintesi del Percorso Formativo \*\*\*\*\* • 39 CFU obbligatori su settori scientifico disciplinari caratterizzanti (I Anno) • 6 CFU a scelta nell'area formazione matematico-statistica (I Anno) • 18 CFU a scelta nell'area della formazione informatica e dell'informazione (I e II Anno) • 6 CFU a scelta nell'area della formazione giuridico, aziendale, linguistica e sociale (I anno) • 12 CFU a scelta nell'area Attività formative affini o integrative (II Anno) • 12 CFU a scelta libera dello studente (I o II Anno) • 3 CFU su Altre conoscenze utili per l'inserimento nel mondo del lavoro (AAF) (II anno) • 24 CFU per la prova finale (II Anno) Le conoscenze raggiunte sono valutate tramite: prove intermedie, discussione di lavori di gruppo, elaborati redatti singolarmente dai discenti ed esami di tipo tradizionale. Piani di studio \*\*\*\*\* Dopo l'immatricolazione, tutti gli studenti devono presentare il proprio piano di studi individuale tramite "Percorsi formativi/GOMP/INFOSTUD". Si precisa che, ad eccezione degli insegnamenti obbligatori, non è possibile verbalizzare un esame non previsto nel piano di studi e il piano di studi deve essere approvato dal docente preposto dal CdS prima di sostenere gli esami. Inoltre, secondo il regolamento di Ateneo, lo studente del 1° anno può sostenere solo gli esami selezionati per il 1° anno, mentre non può sostenere alcun esame previsto per il 2° anno. In particolare, gli studenti del 1° anno non potranno verbalizzare alcun esame che non sia stato previsto per il 1° semestre del 1° anno, prima di un piano di studi sia stato presentato e approvato, ad eccezione degli insegnamenti indicati come "obbligatori". L'elenco degli esami (es. Manifesto degli Studi) a disposizione di uno studente si riferisce sempre all'anno accademico di prima immatricolazione, anche in caso di modifiche negli anni successivi. Il piano di studi degli studenti che si immatricolano dopo aver trasferito la propria carriera da un corso di laurea magistrale in corso sia presso la Sapienza che presso altre Università, verrà compilato automaticamente dalla segreteria subito dopo il riconoscimento dei crediti già acquisiti. Il piano degli studi può essere presentato in determinate finestre temporali e non è possibile compilare per la prima volta o modificare il piano degli studi al di fuori di tali finestre temporali. Finestre temporali e docente di riferimento per i piani di studio saranno indicati sul sito dei Corsi di Studi (<https://corsidilaurea.uniroma1.it/it/>). Caratteristiche della prova finale (24 CFU)

\*\*\*\*\* La prova finale potrà essere inerente ad un'attività progettuale, di ricerca, metodologica o di tirocinio, presso una struttura industriale, istituzioni pubbliche di ricerca o presso i laboratori stessi dell'università. L'esame finale di laurea consiste nella presentazione e discussione di un progetto con caratteri di originalità e di una relazione supervisionata da un docente di riferimento. Il lavoro svolto dovrà dimostrare che lo studente ha raggiunto una padronanza delle metodologie di Data Science e/o della loro applicazione in un settore specifico a un livello di competenza in linea con le esigenze imposte dai processi di innovazione tecnologica. La prova finale sarà impostata in maniera tale da costituire una credenziale importante per l'inserimento del laureato nel tessuto lavorativo. La prova finale si svolgerà mediante una discussione da parte del candidato dinanzi a una Commissione, la cui composizione è stabilita dagli appositi regolamenti di Ateneo. Le Commissioni giudicatrici per la prova finale esprimono la loro votazione in centodecimi e possono, all'unanimità, aggiungere la lode al massimo dei voti. Sbocchi occupazionali e professionali previsti per i laureati \*\*\*\*\* I profili professionali individuati sono quelli di Data Scientist, Open Data Manager, Data Intelligence Professional, Big Data Infrastructure Professional. Le funzioni professionali svolte dai laureati in Data Science sono quelle di analizzare, presentare e prevedere le tendenze fondamentali nei flussi di dati, identificare gli strumenti software necessari per l'elaborazione di grandi moli di dati, coordinare la raccolta e la pubblicazione di open data nel settore pubblico e privato, coordinare gruppi di programmatori e progettare nuove classi di servizi basati sui big data, integrare le metodologie di data science all'interno dei processi organizzativi e di strategie di mercato delle aziende, gestire le più importanti infrastrutture software, hardware e di rete per i big data. Le principali competenze associate alla funzione sono quelle di analisi statistica dei dati, comprensione delle infrastrutture software, data management e data mining, comprensione dei flussi di dati e dei loro formati, management e business analytics, progettazione e analisi di architetture hardware, software e di rete per i big data, competenze giuridiche-economiche nel campo dell'ICT. Gli sbocchi professionali si troveranno all'interno di grandi, medie e piccole aziende, pubblica amministrazione, amministrazioni locali, enti di ricerca pubblici e privati, istituti di analisi economico-sociale ed enti no-profit. Passaggi e Trasferimenti \*\*\*\*\* Gli studenti che intendono trasferirsi al Corso di Laurea Magistrale in Data Science devono presentare domanda presso la Segreteria Studenti delle Facoltà di Ingegneria

dell'Informazione, Informatica e Statistica (c/o città universitaria). La domanda deve essere redatta secondo le modalità previste dalla Segreteria Studenti Amministrativa e dovrà comunque prevedere l'elenco degli esami superati per i quali si richiede il riconoscimento. Il Consiglio di Corso di Studi delibererà gli esami riconosciuti e i crediti attribuiti e provvederà, inoltre, di concerto con lo studente, alla definizione del percorso formativo che, nel rispetto dell'ordinamento didattico e dei contenuti formativi del Corso di Laurea Magistrale in Data Science, potrà tenere conto del percorso già svolto. Norme relative alle iscrizioni ad anni successivi per studenti provenienti da altro corso di laurea o altro ateneo. Per iscriversi al secondo anno del corso di studi lo studente proveniente da altro corso di laurea o altro ateneo deve aver acquisito almeno 27 crediti. Non sono previsti obblighi di frequenza se non per attività di laboratorio o altre attività pratiche o formative. Tempo parziale \*\*\*\*\* I termini e le modalità per la richiesta del regime a tempo parziale, nonché le relative norme, sono stabilite nel Regolamento per gli studenti e le studentesse dei corsi di laurea e laurea magistrale di Ateneo, consultabile sul sito web della Sapienza. E' possibile scegliere un numero di CFU annui che va da 18 a 45. Periodi di studio all'estero \*\*\*\*\* I corsi seguiti nelle Università Europee o estere, con le quali la Facoltà di Ingegneria dell'Informazione, Informatica, Statistica, ha in vigore accordi, progetti e/o convenzioni, vengono riconosciuti secondo le modalità previste dagli accordi e validati dal Responsabile Accademico della Mobilità (RAM). Gli studenti possono, previa autorizzazione del Consiglio di Corso di Studi, svolgere un periodo di studio all'estero nell'ambito del progetto LLP Erasmus. In conformità con il Regolamento didattico di Ateneo nel caso di studi, esami e titoli accademici conseguiti all'estero, il Consiglio di Corso di Studi esamina di volta in volta il programma ai fini dell'attribuzione dei crediti nei corrispondenti settori scientifici disciplinari. Info generali \*\*\*\*\* Programmi e testi d'esame: il programma dei corsi e i testi d'esame sono consultabili sul sito internet: <https://corsidilaurea.uniroma1.it/> Servizi di tutorato I ICdS identifica i docenti che svolgono attività di tutorato e orientamento, secondo le modalità e gli orari indicati sul sito del Corso di Laurea. Inoltre il Corso di Laurea si avvale dei servizi di tutorato messi a disposizione dalla Facoltà, compatibilmente alle risorse economiche, utilizzando anche appositi contratti integrativi. Valutazione della qualità Il Corso di Laurea, in collaborazione con la Facoltà, effettua la rilevazione dell'opinione degli studenti frequentanti per tutti i corsi di insegnamento tenuti. Il sistema di rilevazione è integrato con un percorso qualità la cui responsabilità è affidata al gruppo di auto-valutazione, docenti, studenti e personale del corso di studio. I risultati delle rilevazioni e delle analisi del gruppo di auto-valutazione sono utilizzati per effettuare azioni di miglioramento delle attività formative.

# Assicurazione qualità

## Consultazioni iniziali con le parti interessate

Numerose organizzazioni delle professioni, dei servizi e dell'industria sono state consultate a proposito della proposta di istituire una laurea magistrale in Data Science nella Classe di Laurea LM-DATA presso Sapienza. L'incontro ha avuto luogo in data 20 Settembre 2022 alle ore 10:00 presso l'Aula Magna del Dipartimento di Ingegneria Informatica, Automatica e Gestionale "Antonio Ruberti" della Sapienza, Via Ariosto 25. Sono state consultate grandi, piccole e medie imprese private e pubbliche quali Advant, Amarena SRL, BiP, Centai, Cineca, Comune di Roma Casa delle Tecnologie Emergenti, Crisma Italia, Enel, Engineering, Fondazione Marco Vigorelli, ISMAR-CNR, ISTAT, KPMG, Leonardo, NTT DATA, Poste Italiane, Proceedit.com, Procter & Gamble, Reply, S3K, Telecom/TIM, Unicredit. Tutte le realtà dei servizi, dell'industria e delle professioni intervenute hanno sottolineato l'importanza per la loro realtà lavorativa della figura professionale del data scientist. Tutte hanno indicato la forte richiesta di personale qualificato in grado sia di analizzare le grandi moli di dati e scegliere i sistemi informatici più idonei per la loro conservazione ed elaborazione, sia di dialogare ad ampio spettro con esperti dei vari campi di applicazione. La consultazione ha rilevato positivamente come l'approccio interdisciplinare del corso di studio sia un aspetto fondamentale per la preparazione dei profili professionali richiesti dal mondo del lavoro. Le aziende e gli enti intervenuti hanno tutti apprezzato la proposta di impianto del progetto che vede coinvolti i quattro dipartimenti della Facoltà di Ingegneria dell'Informazione, Informatica e Statistica della Sapienza. Gli intervenuti hanno anche evidenziato come la crescente mole di informazioni generate durante i processi industriali e dei servizi sia un asset strategico della nostra società che richiede figure professionali atte alla loro valorizzazione. Gli enti pubblici hanno sottolineato come l'avvento degli "open data" costituisca uno strumento fondamentale per il processo decisionale politico e amministrativo e una crescente partecipazione dei cittadini alla vita pubblica. Si è sottolineata anche l'importanza della figura di Data Scientist negli enti no profit. Durante la consultazione è stata resa evidente l'eccezionale portata del processo storico in atto che richiede quindi l'urgente formazione di professionisti in grado di analizzare, comprendere e presentare al meglio dati di varia natura in ambiti molto differenti del mondo delle professioni e dei servizi. I verbali dell'incontro sono disponibili al link: <https://drive.google.com/file/d/175MWcYs5yQcqcUuA2MESo2AsTtlcDB5d/view> raggiungibile dal sito web del Dipartimento Dipartimento di Ingegneria Informatica, Automatica e Gestionale "Antonio Ruberti" <https://www.diag.uniroma1.it/corsi-di-studio>

## Consultazioni successive con le parti interessate

La laurea magistrale in Data Science prevede un organo di consultazione permanente con le professioni, i servizi e l'industria (Comitato di indirizzo), che svolge i seguenti compiti: - il monitoraggio dell'aggiornamento del percorso formativo in considerazione delle tendenze in atto nel mercato del lavoro e dell'evoluzione delle esigenze di formazione; - la promozione dello sviluppo di collaborazioni riguardanti le attività di tirocinio, l'orientamento e il supporto ai laureati per il loro ingresso nell'attività lavorativa; - la valutazione dell'efficacia degli esiti occupazionali. Tutte le realtà intervenute nella consultazione effettuata in sede di istituzione del CdS e altre realtà dell'industria e delle professioni si sono rese disponibili a partecipare a tale organo di consultazione. L'interazione con gli stakeholder comprende anche l'organizzazione di eventi, seminari e workshop su temi di attualità, al fine di rendere più pregnanti le attività formative e di sviluppo delle conoscenze, delle competenze e del profilo culturale complessivo degli studenti. Il comitato d'indirizzo è composto da: - Prof. Paolo Di Lorenzo, Prof. Fabrizio Silvestri, Prof. Pierpaolo Brutti, Prof.ssa Veronica Piccialli, Prof. Aris Anagnostopoulos, Prof. Sergio Barbarossa, Prof. Simone Scardapane, Prof. Fabio Galasso; e le seguenti AZIENDE: - Alessandra Bichiri Advant - Romagnoli Raniero Almaxwave - Giuseppe Ricciuti Amarena - Signorino Giuseppe BIP Group - Francesco Bonchi CENTAI - DONATELLA SFORZINI Cineca - Rufini Flavia ENEL - Marco Breda Engineering - Claudia Quadrino Gruppo Olidata - Piero Geraci Humanativa - Paolo Tamagnini KNIME - Cocca Katia P&G - Fabio Luciani S3K - Paolo Barboni Target Reply - Sebastien Bratiars Translated - Ilaria Bordino UniCredit L'incontro tra studenti, industria e docenti avviene nell'ambito dell'Industrial Liaison Program (ILP). Tale programma permette continui contatti con i partner industriali (italiani e multinazionali), istituti di ricerca e dei servizi al fine di individuare ambiti di interesse comune finalizzati all'attivazione di progetti di stage, tesi di laurea magistrale, Internship, e orientamento per il mercato del lavoro. I dettagli degli incontri verranno pubblicati sulla pagina web istituzionale del catalogo corsi di studio della Sapienza (<https://corsidilaurea.uniroma1.it/>) dedicata al corso in Data Science.

## **Organizzazione e responsabilità della AQ del Cds**

Il Sistema di Assicurazione Qualità (AQ) di Sapienza è descritto diffusamente nelle Pagine Web del Team Qualità consultabili all'indirizzo <https://www.uniroma1.it/pagina/team-qualita>. Nelle Pagine Web vengono descritti il percorso decennale sviluppato dall'Ateneo per la costruzione dell'Assicurazione Qualità Sapienza, il modello organizzativo adottato, gli attori dell'AQ (Team Qualità, Comitati di Monitoraggio, Commissioni Paritetiche Docenti-Studenti, Commissioni Qualità dei Corsi di Studio), i Gruppi di Lavoro attivi, le principali attività sviluppate, la documentazione predisposta per la gestione dei processi e delle attività di Assicurazione della Qualità nella Didattica, nella Ricerca e nella Terza Missione. Le Pagine Web rappresentano inoltre la piattaforma di comunicazione e di messa a disposizione dei dati di riferimento per le attività di Riesame, di stesura delle relazioni delle Commissioni Paritetiche Docenti-Studenti e dei Comitati di Monitoraggio e per la compilazione delle Schede SUA-Didattica e SUA-Ricerca. Ciascun Corso di Studio e ciascun Dipartimento ha poi facoltà di declinare il Modello di Assicurazione Qualità Sapienza definito nelle Pagine Web del Team Qualità nell'Assicurazione Qualità del CdS/Dipartimento mutuandolo ed adattandolo alle proprie specificità organizzative pur nel rispetto dei modelli e delle procedure definite dall'Anvur e dal Team Qualità. Le Pagine Web di CdS/Dipartimento rappresentano, unitamente alle Schede SUA-Didattica e SUA-Ricerca, gli strumenti di comunicazione delle modalità di attuazione del Sistema di Assicurazione Qualità a livello di CdS/Dipartimento.