



Principali informazioni sull'insegnamento

Denominazione dell'insegnamento	Apprendimento Automatico
Corso di studio	Data Science
Anno Accademico	2023/24
CFU	6 Crediti formativi universitari
Settore Scientifico Disciplinare	ING-INF/05
Lingua di erogazione	Italiano
Anno di corso	Primo
Periodo di erogazione	2° semestre
Obbligo di frequenza	La frequenza è fortemente raccomandata
Sito web del corso di studio	https://www.uniba.it/it/ricerca/dipartimenti/informatica/didattica/corsi-di-laurea/data-science/data-science

Docente	
Nome e cognome	Nicola Fanizzi
Indirizzo mail	nicola.fanizzi@uniba.it
Telefono	080 544 2246
Sede	Dipartimento di Informatica, Via Orabona 4, 70125, Bari. Stanza n.522, 5° piano.
Sede virtuale	Piattaforma e-learning UNIBA - https://elearning.uniba.it/
Sito web del docente	https://www.uniba.it/it/docenti/fanizzi-nicola
Ricevimento (giorni, orari e modalità, es. su appuntamento)	Martedì, di norma 9:00 - 11:00 previo appuntamento. Orario modificato in caso di sovrapposizione con gli orari di lezione (cfr. avvisi sul sito-docente)

Syllabus	
Obiettivi formativi	L'insegnamento si propone di fornire solide competenze teorico-pratiche nell'ambito dell'Apprendimento Automatico e specificamente capacità di svolgimento di compiti e risoluzione di problemi complessi attraverso metodi per l'apprendimento di modelli predittivi. Capacità di astrazione e, al contempo, di applicazione dei modelli a specifici problemi, nonché la valutazione quantitativa delle loro implementazioni, rappresentano obiettivi non meno importanti dell'insegnamento.
Prerequisiti	<i>Competenze di base (acquisite nei CdS triennali e/o nei corsi del I semestre) relative a:</i> <ul style="list-style-type: none">• <i>Probabilità e Statistica (ins. di Calcolo delle probabilità e statistica, nelle LT)</i>• <i>Algebra Lineare e Analisi Matematica (ins. di Fondamenti di Matematica per la DS)</i>• <i>Programmazione (in Python) (cfr. ins. di Fondamenti di Programmazione per la DS)</i>



<p>Contenuti di insegnamento (Programma)</p>	<ol style="list-style-type: none"> 1. Introduzione 2. Apprendimento supervisionato 3. Teoria delle Decisioni Bayesiana 4. Metodi/Modelli Parametrici 5. Metodi/Modelli Multivariati 6. Riduzione della Dimensionalità: Feature Selection & Extraction 7. Misture / Modelli Semiparametrici, Clustering 8. Metodi/Modelli Non-parametrici 9. Alberi e Regole 10. Discriminazione Lineare: algoritmi basati su gradiente 11. Percettroni Multilayer: Reti Neurali e backpropagation, autoencoder 12. [...] 13. Modelli Locali: approccio competitivo 14. Kernel Machine: SVM per classificazione, regressione, outlier detection 15. Modelli Grafici: Modelli generativi, Inferenze causali e abduitive 16. Hidden Markov Model: problemi fondamentali 17. Approccio Bayesiano 18. Combinazione di Modelli: ECOC, Bagging, Boosting, Stacking, Cascading,... 19. [...] 20. Progettazione e Analisi di Esperimenti: valutazione comparativa di modelli <p><i>Per ciascun argomento: 3 ore circa. la numerazione ricalca quella dei cap. del libro di testo. Per maggiori dettagli si vedano anche i siti dell'insegnamento.</i></p>		
<p>Testi di riferimento</p>	<p>Testo base:</p> <ul style="list-style-type: none"> • E. Alpaydin: Introduction to Machine Learning. MIT Press. 4E <p>Manuali consigliati per le esercitazioni:</p> <ul style="list-style-type: none"> • A. Gèron: Hands-on machine learning with Scikit-Learn. 2e, O'Reilly • Raschka & Mirjalili: Python Machine Learning. 3e. Packt <p>Gli studenti che lo desiderano possono ottenere i testi in prestito dalla Biblioteca. Può convenire verificarne la disponibilità mediante il Sistema Bibliotecario di Ateneo https://opac.uniba.it/easyweb/w8018/index.php? e contattare la biblioteca per concordare il prestito.</p>		
<p>Note ai testi di riferimento</p>	<p>Testo base (4^a ed.): capp. 1-11,13-18,20</p> <p>Le sezioni del programma sono numerate in modo da ritrovare facilmente il contenuto nei capitoli corrispondenti del testo di riferimento (4^a ed). Inoltre, sezioni e sottosezioni di interesse nel capitolo sono facilmente ritrovabili essendo indicati i rispettivi titoli (in italiano).</p> <p>Le slide utilizzate dal docente nelle lezioni/esercitazioni e altro materiale viene messo a disposizione presso i siti dell'insegnamento su ADA e Sharepoint.</p> <ul style="list-style-type: none"> • La parte pratica della prova d'esame prevede un esercizio che ricalca quelli svolti nel corso delle esercitazioni. 		
<p>Organizzazione della didattica</p>			
<p>Ore</p>			
<p>Totali</p>	<p>Didattica frontale</p>	<p>Pratica (esercitazioni in laboratorio)</p>	<p>Studio individuale</p>
<p>225 ore</p>	<p>56 ore</p>	<p>30 ore</p>	<p>139 ore</p>
<p>CFU</p>			
<p>9 CFU</p>	<p>7 CFU</p>	<p>2 CFU</p>	



Metodi didattici	
	<p>Lezioni frontali con l'ausilio di slide, rese disponibili in anticipo.</p> <p>Esercitazioni pratiche in laboratorio (3 ore ciascuna) da svolgere in autonomia o in gruppo con l'utilizzo di ambienti per la programmazione in Python sul proprio dispositivo (BYOD), tipicamente notebook Jupyter e librerie-base ricompresi nella suite Anaconda installata nei laboratori.</p>

Risultati di apprendimento previsti	
Conoscenza e capacità di comprensione	<p>Lo studente dovrà acquisire competenze che, facendo leva sulle conoscenze acquisite nel I ciclo di studi, le estendano in modo da arrivare a una piena consapevolezza sotto il profilo teorico delle possibilità e dei limiti dei modelli e delle metodologie della disciplina, rendendolo capace di concepire autonomamente soluzioni appropriate ad affrontare specifici problemi e che possano risultare originali anche rispetto al contesto di ricerca.</p>
Conoscenza e capacità di comprensione applicate	<p>Lo studente sarà capace di applicare le conoscenze teoriche acquisite, avendo imparato a comprendere e proporre soluzioni a problemi relativi anche a tematiche nuove o non familiari, inserite in contesti più ampi, comprendenti anche altre discipline non strettamente connesse al settore di studio.</p>
Competenze trasversali	<p>Autonomia di giudizio Lo studente acquisirà la capacità di integrare le conoscenze, gestire la complessità, e formulare giudizi anche a fronte di dati limitati o incompleti. Sarà altresì in grado di riflettere sulle responsabilità sociali ed etiche legate all'applicazione delle conoscenze e dei giudizi espressi.</p> <p>Abilità comunicative Lo studente avrà imparato a comunicare in modo chiaro e privo di ambiguità conoscenze e idee-guida sottese alle stesse, nonché opinioni e conclusioni personali, sia a specialisti della materia sia ad altri interlocutori esperti in discipline diverse.</p> <p>Capacità di apprendere in modo autonomo Lo studente svilupperà capacità di apprendimento che gli consentano di continuare ad approfondire lo studio in modo indipendente seguendo propri interessi e inclinazioni.</p>

Valutazione	
Modalità di verifica dell'apprendimento	<p>Prova unica articolata in due momenti successivi (ciascuno della durata di circa 30') che mira ad accertare l'acquisizione di competenze</p> <ul style="list-style-type: none">• pratiche: esercizio (in lab. o su proprio dispositivo) che dimostri la capacità di costruire un workflow per un problema dato, dall'acquisizione del dataset alla valutazione finale dei modelli appresi;• teoriche: colloquio che, partendo dalla discussione dei risultati della prima parte, si estende ad altri aspetti tecnici riguardanti problematiche generali e/o i diversi task, i modelli e le metodologie nel programma dell'insegnamento. <p>Valutazione in trentesimi ottenuta approssimativamente come media ponderata (in proporzione alla tipologia dei cfu) del grado di acquisizione delle competenze accertato nelle due parti.</p>



<p>Criteri di valutazione</p>	<p>Conoscenza e capacità di comprensione:</p> <ul style="list-style-type: none">• Capacità di enunciare con rigore definizioni e teoremi della teoria dell'apprendimento automatico;• Capacità di comprendere e descrivere con precisione e chiarezza le diverse tipologie di problemi di apprendimento, le idee-base a supporto dei diversi modelli (esplorativi o predittivi), e degli algoritmi per il loro apprendimento;• Comprensione della metodologia per l'analisi delle prestazioni dei modelli. <p>Conoscenza e capacità di comprensione applicate:</p> <ul style="list-style-type: none">• Capacità di applicare i risultati teorici appresi;• Livello di competenza nell'uso di ambienti di programmazione per la Data Science nella creazione/trasformazione dei dataset, l'implementazione delle soluzioni;• Comprensione della metodologia per l'analisi delle prestazioni dei modelli. Capacità nel progetto e nell'esecuzione di test statistici delle soluzioni realizzate e nella discussione dei relativi risultati. <p>Autonomia di giudizio:</p> <ul style="list-style-type: none">• Grado di competenza nell'integrazione delle conoscenze, nella scelta di modelli e tecniche appropriate ai diversi problemi, nella valutazione di modelli implementati anche rispetto ad altre soluzioni disponibili <p>Abilità comunicative:</p> <ul style="list-style-type: none">• Grado di competenza nell'acquisizione della terminologia della disciplina e capacità di interlocuzione con esperti ad es. ai fini del confronto di modelli diversi o della discussione di risultati sperimentali <p>Capacità di apprendere:</p> <ul style="list-style-type: none">• Grado di competenza nell'ideazione di possibili <i>estensioni</i> delle tecniche / dei modelli appresi anche in preparazione a insegnamenti più specifici previsti nel semestre successivo (segnatamente <i>Deep Learning</i>)
<p>Criteri di misurazione dell'apprendimento e di attribuzione del voto finale</p>	<p>Parte pratica: p1 (> 50%) Parte teorica: p2 (> 50%) Voto finale: $(2/9) * p1 + (7/9) * p2$ (in trentesimi)</p>



Altro

Si suggerisce agli studenti di affidarsi esclusivamente alle informazioni/comunicazioni fornite sui siti ufficiali del Dipartimento di Informatica, ovvero sui gruppi social solo se costituiti e amministrati esclusivamente dai docenti dei relativi insegnamenti:

- <https://www.uniba.it/it/ricerca/dipartimenti/informatica/didattica/corsi-di-laurea/corsi-di-laurea>
- <https://www.uniba.it/it/ricerca/dipartimenti/informatica>
- <https://elearning.uniba.it/>

I programmi degli insegnamenti sono disponibili qui su

- <https://elearning.uniba.it/>

Le informazioni che tutti gli studenti dovrebbero conoscere sono scritte nei Regolamenti didattici e manifesti degli studi disponibili nel sito:

- <https://www.uniba.it/it/ricerca/dipartimenti/informatica/didattica/corsi-di-laurea/corsi-di-laurea>

Si suggerisce agli studenti di diffidare delle informazioni e dei materiali circolanti su siti o gruppi social non ufficiali, poiché spesso sono risultati non affidabili, non corretti o incompleti. Per ogni dubbio, chiedere un incontro al docente secondo le modalità previste per il ricevimento.

Sito ADA 23/24: presso <https://elearning.uniba.it/course/view.php?id=2077>

Sito Sharepoint legacy: <https://unibari.sharepoint.com/sites/AA>



Basic Information on the Course

Course Title	Apprendimento Automatico
Degree	Data Science
Academic Year	2023/24
ECTS European Credit Transfer and Accumulation System (ECTS), in Italian Crediti Formativi Universitari (CFU)	9 CFU (each CFU corresponds to 25 hours (h) of student's time); CFU are of type T1, T2 or T3 T1 = 8 h lecture + 17 h individual study T2 = 15 h practice + 10 h individual study T3 = 25 h individual study
SSD	ING-INF/05
Language	Italian
Year	First
Semester	2nd [Spring]
Attendance	Highly recommended
Degree website	https://www.uniba.it/it/ricerca/dipartimenti/informatica/didattica/corsi-di-laurea/data-science/data-science

Instructor

Name and surname	Nicola Fanizzi
Email address	nicola.fanizzi@uniba.it
Telephone number	080 544 2246
Office	Dipartimento di Informatica, Via Orabona 4, 70125, Bari. room n.522, 5th floor
E-learning platform	https://elearning.uniba.it
Instructor's website	https://www.uniba.it/it/docenti/fanizzi-nicola
Office hours	Tue, 9-11 by appointment to be arranged in advance. (modifications notified through the course website)

Syllabus

Educational goals	The course aims at providing sound theoretical-practical competences in the field of Machine Learning and specifically skills in performing tasks and solving complex problems through methods for learning predictive models. Abstraction skills and, at the same time, the application of models to specific problems, as well as the quantitative evaluation of their implementations, are equally important teaching objectives.
Prerequisites	<i>Basic skills (acquired in BSc courses and/or first-semester courses) related to:</i> <ul style="list-style-type: none">• <i>Probability and Statistics</i> (Calcolo delle probabilità e statistica course)• <i>Linear Algebra and Calculus</i> (Fondamenti di Matematica per la DS course)• <i>Python programming</i> (Fondamenti di Programmazione per la DS course)



<p>Course Content (Programme)</p>	<ol style="list-style-type: none"> 1. Introduction 2. Supervised Learning 3. Bayesian Decision Theory 4. Parametric Models 5. Multivariate Models 6. Dimensionality reduction: Feature Selection & Extraction 7. Clustering: Semiparametric Mixture Models 8. Nonparametric Models 9. Trees and Rules 10. Linear Discrimination: gradient based algorithms 11. Multilayer Perceptrons: Neural Networks, backpropagation, autoencoders 12. [...] 13. Local Models: competitive approach 14. Kernel Machines: SVM for classification, regression, outlier detection 15. Graphical Models: generative models, causal and abductive inference 16. Hidden Markov Models: basic problems 17. Bayesian Approach 18. Composite Models: ECOC, Bagging, Boosting, Stacking, Cascading,... 19. [...] 20. Experimental Design and Analysis: comparative evaluation of models <p><i>For each topic about 3 hours. The numbering follows that of the chapters in the textbook. See also the course websites for more details.</i></p>		
<p>Reference textbooks</p>	<p>Base textbook:</p> <ul style="list-style-type: none"> • E. Alpaydin: Introduction to Machine Learning. MIT Press. 4E <p>Further handbooks for exercises:</p> <ul style="list-style-type: none"> • A. Gèron: Hands-on machine learning with Scikit-Learn. 2e, O'Reilly • Raschka & Mirjalili: Python Machine Learning. 3e. Packt <p>Students may borrow texts from the library, checking their availability via the University Library System https://opac.uniba.it/easyweb/w8018/index.php? and contact the library to arrange the loan.</p>		
<p>Notes on the textbooks</p>	<p>Textbook (4th ed.): chapters 1-11,13-18,20</p> <p>The sections of the programme are numbered so that the content can be easily found in the corresponding chapters of the reference text (4th ed). In addition, sections and subsections of interest in the chapter are easily found as their respective titles are indicated.</p> <p>The slides used by the lecturer in the lectures/lab lessons and other materials are made available at the teaching sites on ADA and Sharepoint.</p> <ul style="list-style-type: none"> • The practical part of the examination involves an exercise that replicates those carried out during the lab training lessons 		
<p>Activity Organization</p>			
<p>Hour</p>			
<p>Totals</p>	<p>Front lectures</p>	<p>Practice (lab training)</p>	<p>Individual study</p>
<p>225 hrs</p>	<p>56 hrs</p>	<p>30 hrs</p>	<p>139 hrs</p>
<p>ETCS</p>			
<p>9 ETCS</p>	<p>7 ETCS</p>	<p>2 ETCS</p>	



Teaching Modes	
	Lectures delivered with the aid of slides, made available in advance. Practical lab exercises (3 hours) to be carried out independently or in groups using environments for programming in Python on one's own device (BYOD), typically Jupyter notebooks and base libraries included in the Anaconda suite installed in the labs.

Expected Learning Outcomes	
Knowledge and comprehension skills	The student is expected to develop skills that, building on the knowledge acquired in the first cycle of studies, extend it to a full theoretical awareness of the possibilities and limits of the discipline's models and methodologies, enabling him or her to independently devise solutions that are appropriate to address specific problems and that can also be original with respect to the research context.
Applied knowledge and comprehension skills	The student will be able to apply the theoretical knowledge acquired, having learnt to understand and propose solutions to problems also relating to new or unfamiliar topics, set in broader contexts, including other disciplines not strictly related to the field of study.
Further skills	Autonomy of judgment The student will acquire the ability to integrate knowledge, handle complexity, and make judgments even in the face of limited or incomplete data. He or she will also be able to reflect on the social and ethical responsibilities involved in applying the knowledge and judgments made. Communication skills The student will have learned to clearly and unambiguously communicate knowledge and underlying ideas-guidelines, as well as personal opinions and conclusions, to both subject specialists and other expert counterparts in different disciplines. Ability to learn autonomously The student will develop learning skills that enable him or her to continue to study autonomously following his or her own interests and inclinations.

Evaluation	
Assessment of learning outcomes	Single test divided into two successive moments (each lasting about 20-30') aimed at ascertaining the acquisition of <ul style="list-style-type: none">• <i>practical skills</i>: exercise (in the lab. or on one's own device) demonstrating the ability to construct a workflow for a given problem, from the acquisition of the dataset to the final evaluation of the learned models;• <i>theoretical skills</i>: colloquium that, starting from the discussion of the results of the first part, extends to other technical aspects concerning general problems and/or the different tasks, models and methodologies in the teaching program. Evaluation in thirtieths obtained approximately as a weighted average (in proportion to the type of credits) of the degree of acquisition of skills ascertained in the two parts.



<p>Evaluation Criteria</p>	<p>Knowledge and comprehension skills:</p> <ul style="list-style-type: none"> • Ability to rigorously state definitions and theorems of machine learning theory; • Ability to accurately and clearly understand and describe the different types of learning problems, the basic ideas supporting different models (exploratory or predictive), and the algorithms for learning them; • Understanding of the methodology for analyzing the performance of models. <p>Applied knowledge and understanding skills:</p> <ul style="list-style-type: none"> • Ability to apply learned theoretical results; • Level of proficiency in the use of Data Science programming environments in creating/transforming datasets, implementing solutions; • Understanding of methodology for model performance analysis. Proficiency in the design and execution of statistical tests of implemented solutions and discussion of the results. <p>Autonomy of judgment:</p> <ul style="list-style-type: none"> • Level of proficiency in integrating knowledge, choosing models and techniques appropriate to different problems, evaluating implemented models even in comparison with other available solutions <p>Communication skills:</p> <ul style="list-style-type: none"> • Level of competence in acquiring the terminology of the discipline and ability to engage with experts e.g., for the purpose of comparing different models or discussing experimental results <p>Ability to learn:</p> <ul style="list-style-type: none"> • Level of competence in devising possible extensions of techniques/models learned also in preparation for more specific teachings scheduled for the following semester (notably <i>Deep Learning</i>)
<p>Criteria for assessing learning outcomes and final grading</p>	<p>Part 1, practice: p1 (> 50%) Part 2, theory: p2 (> 50%)</p> <p>Final grade: $(2/9) * p1 + (7/9) * p2$ (in thirtieths)</p>
<p>Further Information</p>	<p>It is suggested that students rely exclusively on the information/communication provided on the official websites of the Department of Computer Science, or on social groups only if they are set up and maintained exclusively by the instructors of the respective courses:</p> <ul style="list-style-type: none"> • https://www.uniba.it/it/ricerca/dipartimenti/informatica/didattica/corsi-di-laurea/corsi-di-laurea • https://www.uniba.it/it/ricerca/dipartimenti/informatica • https://elearning.uniba.it/ <p>Course syllabi are available at the site:</p> <ul style="list-style-type: none"> • https://elearning.uniba.it <p>Information that all students are expected to know is written in the Teaching Regulations and Study Manifestos available on the site:</p> <ul style="list-style-type: none"> • https://www.uniba.it/it/ricerca/dipartimenti/informatica/didattica/corsi-di-laurea/corsi-di-laurea <p>Students are suggested to be wary of information and materials circulating on unofficial sites or social groups, as they are often found to be unreliable, incorrect or incomplete. If you have any doubts, ask for a meeting with the lecturer in accordance with the arrangements for the office hours.</p> <hr/> <p>ADA 23/24 website: https://elearning.uniba.it/course/view.php?id=2077 Sharepoint legacy website: https://unibari.sharepoint.com/sites/AA</p>