

Principali informazioni sull'insegnamento	
Titolo insegnamento	Apprendimento Automatico
Corso di studio	LM Data Science
Crediti formativi	9
Denominazione inglese	Machine Learning
Obbligo di frequenza	no
Lingua di erogazione	Italiano

Docente responsabile	Nome Cognome	Indirizzo Mail
	Nicola Fanizzi	nicola.fanizzi@uniba.it

Dettaglio crediti formativi	Ambito disciplinare	SSD	Crediti
	INFORMATICO	ING-INF/05	7 (T1) 2 (T2)

Modalità di erogazione	
Periodo di erogazione	Secondo semestre
Anno di corso	Primo
Modalità di erogazione	Lezioni frontali Esercitazioni guidate (in aula o lab)

Organizzazione della didattica	
Ore totali	225
Ore di corso	86 (56+30)
Ore di studio individuale	139 (119+20)

Calendario	
Inizio attività didattiche	1 marzo 2021
Fine attività didattiche	4 giugno 2021

Syllabus	
Prerequisiti	Competenze di base relative a Probabilità e Statistica, Algebra Lineare e Analisi Matematica (funzioni di più variabili); Programmazione (in Python)
Risultati di apprendimento previsti (declinare rispetto ai Descrittori di Dublino) (si raccomanda che siano coerenti con i risultati di apprendimento del CdS, compresi i risultati di apprendimento trasversali)	<ul style="list-style-type: none"> • Conoscenza e capacità di comprensione: dei problemi e dei modelli con i relativi algoritmi • Conoscenza e capacità di comprensione applicate: progettazione, implementazione e collaudo di soluzioni basate sui suddetti modelli/algoritmi • Autonomia di giudizio: capacità di confronto di modelli e

	<p>algoritmi; discussione critica dei risultati dei test</p> <ul style="list-style-type: none"> • Abilità comunicative: capacità di relazionare sulla soluzioni proposte anche in termini comparativi • Capacità di apprendere: ideazione e progettazione di opportuni adattamenti dei modelli a nuovi problemi da risolvere
Contenuti di insegnamento	<ul style="list-style-type: none"> • Introduzione; Apprendimento supervisionato; Teoria delle Decisioni Bayesiana; • Metodi Parametrici; Metodi Multivariati; Selezione/Estrazione di Feature*; • Modelli Misti Semiparametrici; • Metodi Non-parametrici; Alberi e Regole*; • Discriminazione Lineare; Modelli Neurali; • Modelli Locali; Kernel Machine; • Modelli Grafici*; Hidden Markov Model*; • Approccio Bayesiano; • Combinazione di Modelli; • Reinforcement Learning*; • Progettazione e Analisi di Esperimenti

Programma	<ol style="list-style-type: none"> 1. Introduzione: Scenario: definizione, problemi, soluzioni, modelli; applicazioni: associazione, classificazione, regressione, apprendimento non supervisionato, apprendimento per rinforzo 2. Apprendimento supervisionato: Classificazione, dimensione VC; PAC learning; rumore; classi molteplici; regressione; Model Selection e generalizzazione; dimensioni degli algoritmi supervisionati 3. Teoria delle Decisioni Bayesiana: Richiami su Probabilità e Inferenza; Classificazione: Regola di Bayes; Loss e Rischio; Rigetto; Funzioni Discriminanti: Utility Theory; Regole di Associazione (cenni): misure di associazione, Apriori 4. Metodi Parametrici: Stima Parametrica, caso univariato; Stima di Massima Verosimiglianza; Valutazione di stimatori: Bias e Varianza; Stimatore di Bayes Predizione, Stimatori MAP, ML e di Bayes; Classificazione; Regressione: dalla Log-Likelihood all'Errore, misure; Complessità dei modelli: Dilemma Bias / Varianza; Model Selection: Reg. Polinomiale, selezione Bayesiana; 5. Metodi Multivariati: Modelli multivariati; stima dei parametri; Stima di valori mancanti; Normale multivariata: bivariata, Naive Bayes; Classificazione multivariata: classificazione parametrica, stima di parametri, matrici di covarianza; Complessità: model selection; Feature discrete; Regressione multivariata 6. Selezione/Estrazione di Feature (cenni): Subset selection; PCA; Feature Embedding, Factor Analysis; SVD; MDS; LDA; CCA; Isomap; LLE, Laplacian Eigenmaps 7. Modelli Misti Semiparametrici: Introduzione: density mixture, classi e cluster; k-Means: quantizzazione, codifica/decodifica, preprocessing; EM; Mixture model con variabili latenti; Caso supervisionato; Clustering spettrale; Modelli gerarchici basati su distanze; Scelta di k; 8. Metodi Non-parametrici: Stima non parametrica della densità: istogrammi, kernel, NN; Dati multivariati; Classificazione non parametrica; Condensed NN; Metodi di classificazione basati su distanza; Outlier; Regressione non
-----------	--

- parametrica: Modelli di *Smoothing*; Scelta dei parametri di *smoothing*
9. **Alberi e Regole** (cenni): Alberi di decisione e regressione, Potatura; Regole: estrazione; Apprendimento diretto di regole dai dati; Alberi multivariati ed estensioni
 10. **Discriminazione Lineare**: Classificazione; Generalizzazioni del modello lineare, Geometria; Separabilità; Discriminazione parametrica; Discesa di gradiente; Discriminazione logistica; Discriminazione tramite Regressione; *Ranking*
 11. **Modelli Neurali**: Reti Neurali; Percettrone; Training; Apprendimento di funzioni booleane; MLP: approssimazione universale; *Backpropagation*: layer nascosti multipli; Procedure di addestramento; Regolazione delle dimensioni della rete: Adattamento strutturale (distruttivo e costruttivo) incrementale; Apprendimento bayesiano; Riduzione della Dimensionalità: Autoencoder, MLP e MDS; Apprendimento Temporale: *Time-Delay Neural Network*, *Recurrent Network*; *Deep Learning* (nozioni-base)
 12. **Modelli Locali**: Apprendimento competitivo: *Online k-Means*, *Winner-take-all network*, *Hebbian Learning**, ART, SOM; *Radial basis function*: rappresentazione locale/distribuita, RBF Network, Regole ed eccezioni; Incorporare conoscenza in forma di regole; *Normalized basis functions*: EM per RBF (EM supervisionato); LVQ; *Mixture of Experts*: cooperativa, competitiva; gerarchica
 13. **Kernel Machine**: Iperpiano separatore ottimale: margine, *Support Vector Machine*; Non-separabilità lineare: Margine *soft*, *Hinge Loss*; *Nu-SVM*; *Kernel Trick* e BF; Kernel vettoriali; *Multiple Kernel Learning*; Caso multiclasse; Regressione; *Ranking*; One-Class; *Large Margin NN*; Apprendimento di distanze; Riduzione della dimensionalità
 14. **Modelli Grafici** (cenni): Modelli Grafici: Regola di Bayes, Indipendenza condizionata; Casi canonici; Inferenza causale e diagnostica, Sfruttamento della struttura locale; Modelli generativi: Classificazione, *Naive Bayes*, Regressione lineare; d-Separazione: blocco; *Belief Propagation*: catene, alberi, polialberi, alberi di giunzione; Grafi non orientati: *Markov Random Fields*, *Factor Graph*; Apprendimento della struttura; *Influence diagrams*
 15. **Hidden Markov Model** (cenni): Processi di Markov discreti: automa stocastico; HMM: sviluppo nel tempo, problemi di base; Valutazione; Ricerca della sequenza di stati: algo di *Viterbi*; Apprendimento dei parametri: *Baum-Welch* (EM); Osservazioni nel continuo: HMM con input; come modello grafico; *Model Selection*
 16. **Approccio Bayesiano**: processi gaussiani, processi di Dirichlet, ristorante cinese e ristorante indiano; Modelli generativi: Stima bayesiana dei parametri di distribuzioni discrete: caso binario ed estensioni; Distribuzioni continue: Gaussiane univariate e multivariate; Stima dei parametri di una funzione: regressione, kernel; scelta della distribuzione a priori; Modelli bayesiani: confronto; Stima bayesiana di *mixture model*; Modellazione bayesiana non-parametrica: processi gaussiani, processi di Dirichlet e ristorante cinese; *Gaussian mixture non-parametrica*; *Latent Dirichlet Allocation*; Processi beta e Buffet Indiano: *Feature Extraction* bayesiana
 17. **Combinazione di Learner**: *Ensemble*; Schemi di combinazione; *Voting*: regole fisse, approccio bayesiano, ECOC; *Boosting*, *Bagging*, MoE; *Stacking*; *Cascading*; Forme di integrazione
 18. **Reinforcement Learning** (cenni): Apprendimento con Critico; *K-armed Bandit*; Elementi: MDP, *policy* e *reward*, ottimizzazione; *Model-Based Learning*: *value / policy iteration*; *Temporal Difference Learning*: Strategie d'esplorazione: *reward* e azioni deterministici e non; *Q-learning*: algoritmi *on-policy*; Sarsa: *TD Learning*; Generalizzazione a problemi di regressione; POMDP: stati parzialmente osservabili
 19. **Progettazione e Analisi di Esperimenti**: Composizione dei dataset, fattori di

	preferenza degli algoritmi; Esperimenti: fattori, risposta e strategie, principi di base e linee guida; <i>Cross-Validation</i> e metodi di <i>Resampling</i> : K-Fold CV, 5×2 CV, <i>Bootstrapping</i> ; Misurazione delle prestazioni: misure, curva ROC, <i>precision + recall</i> ; Stima degli intervalli; Test di Ipotesi: valutazione delle prestazioni; Errore nella classificazione, Test binomiale e approssimazione normale, <i>Paired t Test</i> ; Confronto di classificatori: <i>K-Fold CV Paired t Test</i> ; <i>5×2 CV Paired t Test</i> , <i>5×2 CV Paired F Test</i> ; Confronto multiplo: test ANOVA; Confronto su più Dataset; Test multivariati: a coppie o ANOVA multivariato
Testi di riferimento	<p>Testo base: E. Alpaydin: <i>Introduction to Machine Learning</i>. MIT Press. 3e (4e)</p> <p>Manuale per le esercitazioni: Raschka & Mirjalili: <i>Python Machine Learning</i>. 3e. Packt</p> <p>Altri testi di riferimento: indicati sul sito ADA dell'insegnamento</p>
Note ai testi di riferimento	Alpaydin: capp. 1-19 Raschka & Mirjalili: capp. 1-14
Metodi didattici	Lezioni frontali ed esercitazioni pratiche con l'utilizzo di ambienti per la programmazione Python (cfr. suite Anaconda). Uso delle piattaforme di E-learning per esercizi da svolgere autonomamente.
Metodi di valutazione (indicare almeno la tipologia scritto, orale, altro)	Prova orale con accertamento delle competenze teorico-tecniche acquisite
Criteri di valutazione (per ogni risultato di apprendimento atteso su indicato, descrivere cosa ci si aspetta lo studente conosca o sia in grado di fare e a quale livello al fine di dimostrare che un risultato di apprendimento è stato raggiunto e a quale livello)	<p>Grado di capacità nell'individuare la tipologia del problema da risolvere al fine di scegliere soluzioni adeguate anche tramite opportuni adattamenti.</p> <p>Capacità di discussione su modelli e algoritmi e anche in termini di complessità. Grado di conoscenza della letteratura.</p> <p>Livello di competenza nell'uso di ambienti di programmazione per la <i>Data Science</i> nella creazione/trasformazione dei dataset, l'implementazione delle soluzioni.</p> <p>Grado di capacità nel progetto e nell'esecuzione di test statistici delle soluzioni realizzate e nella discussione dei relativi risultati.</p>
Altro	