

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	24 febbraio 2020
Fine attività didattiche	29 maggio 2020

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 algoritmi; discussione critica dei risultati dei test • 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)
-----------	---

	<p>12. Modelli Locali: Apprendimento competitivo: <i>Online k-Means, Winner-take-all network, Hebbian Learning*</i>, ART, SOM; <i>Radial basis function:</i> rappresentazione locale/distribuita, RBF Network, Regole ed eccezioni; Incorporare conoscenza in forma di regole; <i>Normalized basis functions:</i> EM per RBF (EM supervisionato); LVQ; <i>Mixture of Experts:</i> cooperativa, competitiva; gerarchica</p> <p>13. Kernel Machine: Iperpiano separatore ottimale: margine, <i>Support Vector Machine;</i> Non-separabilità lineare: Margine <i>soft, Hinge Loss; Nu-SVM;</i> <i>Kernel Trick</i> e BF; Kernel vettoriali; <i>Multiple Kernel Learning;</i> Caso multiclasse; Regressione; <i>Ranking;</i> One-Class; <i>Large Margin NN;</i> Apprendimento di distanze; Riduzione della dimensionalità</p> <p>14. Modelli Grafici (cenni): Modelli Grafici: Regola di Bayes, Indipendenza condizionata; Casi canonici; Inferenza causale e diagnostica, Sfruttamento della struttura locale; Modelli generativi: Classificazione, <i>Naive Bayes,</i> Regressione lineare; d-Separazione: blocco; <i>Belief Propagation:</i> catene, alberi, polialberi, alberi di giunzione; Grafi non orientati: <i>Markov Random Fields, Factor Graph;</i> Apprendimento della struttura; <i>Influence diagrams</i></p> <p>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 <i>Viterbi;</i> Apprendimento dei parametri: <i>Baum-Welch</i> (EM); Osservazioni nel continuo: HMM con input; come modello grafico; <i>Model Selection</i></p> <p>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 <i>mixture model;</i> Modellazione bayesiana non-parametrica: processi gaussiani, processi di Dirichlet e ristorante cinese; <i>Gaussian mixture</i> non-parametrica; <i>Latent Dirichlet Allocation;</i> Processi beta e Buffet Indiano: <i>Feature Extraction</i> bayesiana</p> <p>17. Combinazione di Learner: <i>Ensemble;</i> Schemi di combinazione; <i>Voting:</i> regole fisse, approccio bayesiano, ECOC; <i>Boosting, Bagging, MoE; Stacking; Cascading;</i> Forme di integrazione</p> <p>18. Reinforcement Learning (cenni): Apprendimento con Critico; <i>K-armed Bandit;</i> Elementi: MDP, <i>policy</i> e <i>reward,</i> ottimizzazione; <i>Model-Based Learning: value / policy iteration; Temporal Difference Learning:</i> Strategie d'esplorazione: <i>reward</i> e azioni deterministici e non; <i>Q-learning:</i> algoritmi <i>on-policy;</i> Sarsa: <i>TD Learning;</i> Generalizzazione a problemi di regressione; POMDP: stati parzialmente osservabili</p> <p>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, 5x2 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; 5x2 CV Paired t Test, 5x2 CV Paired F Test;</i> Confronto multiplo: test ANOVA; Confronto su più Dataset; Test multivariati: a coppie o ANOVA multivariato</p>
Testi di riferimento	<p>Testo base: E. Alpaydin: Introduction to Machine Learning. MIT Press. 3e (4e)</p> <p>Manuale per le esercitazioni: Raschka & Mirjalili: Python Machine Learning. 3e. Packt</p> <p>Altri testi di riferimento: indicati sul sito ADA dell'insegnamento</p>
Note ai testi di riferimento	<p>Alpaydin: capp. 1-19 Raschka & Mirjalili: capp. 1-14</p>
Metodi didattici	<p>Lezioni frontali ed esercitazioni pratiche con l'utilizzo di ambienti per la programmazione Python (cfr. suite Anaconda).</p> <p>Uso delle piattaforme di E-learning per esercizi da svolgere autonomamente.</p>

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	