Testi del Syllabus

Resp. Did.	**BONONI Alberto**	Matricola:	**004854**

Anno offerta:	**2015/2016**
Insegnamento:	**1006078 - MACHINE LEARNING FOR PATTERN RECOGNITION (1st MODULE)**
Corso di studio:	**5052 - COMMUNICATION ENGINEERING - INGEGNERIA DELLE TELECOMUNICAZIONI**
Anno regolamento:	**2015**
CFU:	**3**
Settore:	**ING-INF/05**
Tipo Attività:	**C - Affine/Integrativa**
Anno corso:	**1**
Periodo:	**Secondo Semestre**
Sede:	**PARMA**

Testi in italiano

Tipo testo	***Testo***
Lingua insegnamento	Inglese
Contenuti	PART 1: Fundamentals (Bononi): Basic probability refresher. Bayesian binary and M-ary classification. MAP and Minimax classifiers. Performance and ROC. Gaussian case and linear discriminant rules. Bayesian estimation (regression). Maximum likelihood, MMSE, MMAE estimators. Linear suboptimal estimators. Supervised learning. Generative versus discriminative approaches. Plug-in learning. Bayesian learning. Minimum empirical risk learning. Nonparametric probability density estimation. linear data reduction for feature extraction. PART 2: Advanced topics and applications (Cagnoni) Support Vector Machines. Classifier evaluation techniques. Unsupervised classification and clustering. - K-means and Isodata algorithms - Self-Organizing Maps - Learning Vector Quantization - Kohonen networks
Testi di riferimento	[1] C. W. Therrien, "Decision, estimation and classification" Wiley, 1989 [2] R. O. Duda, P. E. Hart, D. G. Stork, "Pattern classification", 2nd Ed., Wiley, 2001 [3] D. Barber "Bayesian Reasoning and Machine Learning" Cambridge University Press, 2012. [4] C. M. Bishop "Pattern Recognition and Machine Learning", Springer, 2006. [5] T. Hastie, R. Tibshirani, J. Friedman, "The Elements of Statistical Learning: Data mining, inference, and prediction", Springer, 2008.

Tipo testo	*Testo*
Obiettivi formativi	L'obiettivo del corso è fornire allo studente la capacità di comprendere ed applicare le regole di base dell'apprendimento automatico, e in particolare: - applicare i principali test statistici nella classificazione tra diverse categorie - sintetizzare la struttura del classificatore ottimo e valutarne l'errore di classificazione - applicare i principali metodi di estrazione delle features dai dati - applicare i principali stimatori in uso nel campo dell'apprendimento automatico - applicare i principali algoritmi di clustering nell'apprendimento non supervisionato Le capacità di applicare le conoscenze sopra elencate risultano essere in particolare: - progettare ed analizzare le prestazioni di un classificatore nell'apprendimento automatico - selezionare le feature più appropriate per discriminare le categorie di ingresso - selezionare gli algoritmi di clustering più appropriati nella progettazione di un classificatore non supervisionato.
Prerequisiti	Corsi di base in algebra lineare e teoria della probabilità, quali ad esempio quelli offerti nel corso di laurea triennale corrispondente, sono necessari prerequisiti per questo corso.
Metodi didattici	Didattica frontale 42 ore. Esercitazioni 6 ore. Esercizi assegnati per casa.
Altre informazioni	Ricevimento Bononi: Lunedi' 11:30-13:30 (Sede Scientifica Ingegneria, Pal. 2, I piano, Stanza 2/19T). Cagnoni: su appuntamento(Sede Scientifica Ingegneria, Pal.1, I piano, email cagnoni[AT]ce.unipr.it).
Modalità di verifica dell'apprendimento	Part 1, Bononi: Esame orale, su appuntamento. Al momento dell'iscrizione, contattare il docente all'indirizzo alberto.bononi[AT]unipr.it specificando la data desiderata. L'esame richied la soluzione di alcuni esercizi e la discussione dei dettagli teorici ad essi collegati, pe runa durata di circa un'ora. E' consentito l'uso di un formulario su un foglio A4 per consultazione durante l'esame. Part 2, Cagnoni: Assegnazione di un progetto pratico, i cui risultati dovranno essere presentati in forma di relazione scritta e di presentazione orale.
Programma esteso	Syllabus PART 1: Fundamentals (Bononi) ____ Lec. 1. Introduction - Problem statement and definitions - Examples of machine learning problems - Glossary of equivalent terms in Radar detecton theory, hypothesis testing and machine learning Lec. 2. Probability refresher - Axioms, conditional probability, total probability law, Bayes law, double conditioning, chain rule, independence and conditional independence of events. - Discrete random variables (RV): expectation, conditional expectation. Pairs of RVs. Sum rule. Iterated expectation. Vectors of RVs. An extended example. Lec. 3. Probability refresher - Random vectors: expectation, covariance and its properties, spectral decomposition of

Tipo testo	*Testo*
	covariance matrix, whitening. - Continuous RV. Parallels with discrete RVs. Functions of RVs. Mixed RVs. Continuous random vectors. - Appendix: differentiation rules for vectors and matrices. Lec. 4. - Gaussian RVs and their linear transformations. Mahalanobis distance. Classification: - Bayesian prediction: introduction, loss function, conditional risk, argmin/argmax rules - Bayes classification: introduction Lec. 5. Classification - 0/1 loss -> maximum a posteriori (MAP) classifier. Binary MAP. Decision regions. - Classifier performance. - Likelihood ratio tests and receiver operating curve (ROC) - Minimax rule Lec. 6. Classification - Binary Gaussian classification - Homoscedastic case: linear discriminant analysis - Heteroscedastic case: Bhattacharrya bound - Bayes classification with discrete features - Classification with missing data (composite hypothesis testing) Lec. 7. Estimation - Bayesian estimation: introduction - Quadratic loss: minimum mean square error (MMSE) estimator = regression curve - L1 loss: minimum mean absolute error (MMAE) estimator - 0/1 loss: MAP estimator, and maximum likelihood (ML) in uniform prior. - Regression for vector Gaussian case - ML estimation for Gaussian observations Lec. 8. Estimation - ML for multinomial - Conjugate priors in MAP estimation - Estimation accuracy and ML properties, Cramer Rao bounds. Suboptimal (non Bayesian) estimation: - LMMSE estimation (linear regression) - LMMSE derivation with LDU decomposition Lec. 9. Estimation - LMMSE examples - Generalized linear regression - Example: polynomial regression - Sample LMMSE - Generalized sample LMMSE. Lec. 10. Learning - Supervised learning: introduction - Generative vs discriminatie approaches - Example: logistic model - Plug-in learning ML fitting of logistic model: logistic regression Example: handwritten digit recognition. - Bayesian Learning Lec. 11. Learning: - Empirical risk minimization Nonparametric density estimation: - Parzen window estimator - kNN estimator

Tipo testo	***Testo***
	Lec. 12. linear data reduction - Principal component analysis (PCA) - Fisher linear classifier PART 2: Advanced topics and applications (Cagnoni) ____ --- Linear Discriminant Analysis: - Support Vector Machines --- Classifier evaluation: - generalization and overfitting (Training/validation/test sets) - performance indices, representations curve, confusion matrices - Classification risk: are all errors equally relevant ? --- Unsupervised classification and clustering - K-means and Isodata algorithms - Self-Organizing Maps - Learning Vector Quantization - Kohonen networks

Testi in inglese

Tipo testo	***Testo***
Lingua insegnamento	English
Contenuti	PART 1: Fundamentals (Bononi): Basic probability refresher. Bayesian binary and M-ary classification. MAP and Minimax classifiers. Performance and ROC. Gaussian case and linear discriminant rules. Bayesian estimation (regression). Maximum likelihood, MMSE, MMAE estimators. Linear suboptimal estimators. Supervised learning. Generative versus discriminative approaches. Plug-in learning. Bayesian learning. Minimum empirical risk learning. Nonparametric probability density estimation. linear data reduction for feature extraction. PART 2: Advanced topics and applications (Cagnoni) Support Vector Machines. Classifier evaluation techniques. Unsupervised classification and clustering. - K-means and Isodata algorithms - Self-Organizing Maps - Learning Vector Quantization - Kohonen networks
Testi di riferimento	[1] C. W. Therrien, "Decision, estimation and classification" Wiley, 1989 [2] R. O. Duda, P. E. Hart, D. G. Stork, "Pattern classification", 2nd Ed., Wiley, 2001 [3] D. Barber "Bayesian Reasoning and Machine Learning" Cambridge University Press, 2012. [4] C. M. Bishop "Pattern Recognition and Machine Learning", Springer, 2006. [5] T. Hastie, R. Tibshirani, J. Friedman, "The Elements of Statistical Learning: Data mining, inference, and prediction", Springer, 2008.
Obiettivi formativi	Course Objectives ----------------- The objective of the course is to provide the student with the ability to understand and apply the basic rules of machine learning and, in particular: - to apply the most common statistical tests in classification among

Tipo testo	*Testo*
	different categories - to synthesize the structure of the optimal classifier and analyze its error performance - to apply the most common feature extraction methods from input data - to apply the most common statistical estimators in machine learning - to apply the most common clustering algorithms in unsupervised learning The abilities in applying the above-mentioned knowledge are in particular in the: - design and performance analysis of classifiers in machine learning - selection of the most appropriate features to discriminate input categories - selection of the most appropriate clustering algorithms in the design of unsupervised classifiers
Prerequisiti	Entry-level courses in linear algebra and probability theory, such as those normally offered in the corresponding 3-year Laurea course, are necessary pre-requisites for this course.
Metodi didattici	Classroom teaching, 42 hours. In-class problem solving, 6 hours. Homework regularly assigned.
Altre informazioni	Office Hours Bononi: Monday 11:30-13:30 (Scientific Complex, Building 2, floor 2, Room 2/19T). Cagnoni: by appointment (Scientific Complex, Building 1, floor 2, email cagnoni[AT]ce.unipr.it).
Modalità di verifica dell'apprendimento	Part 1, Bononi: Oral only, to be scheduled on an individual basis. When ready, please contact the instructor by email at alberto.bononi[AT]unipr.it and by specifying the requested date. The exam consists of solving some exercises and explaining theoretical details connected with them, for a total time of about 1 hour. You can bring your summary of important formulas in an A4 sheet to consult if you so wish. Part 2, Cagnoni: A practical project will be assigned, whose results will be presented and discussed by the student both as a written report and as an oral presentation.
Programma esteso	Syllabus PART 1: Fundamentals (Bononi) ______________________________________ Lec. 1. Introduction - Problem statement and definitions - Examples of machine learning problems - Glossary of equivalent terms in Radar detecton theory, hypothesis testing and machine learning Lec. 2. Probability refresher - Axioms, conditional probability, total probability law, Bayes law, double conditioning, chain rule, independence and conditional independence of events. - Discrete random variables (RV): expectation, conditional expectation. Pairs of RVs. Sum rule. Iterated expectation. Vectors of RVs. An extended example. Lec. 3. Probability refresher - Random vectors: expectation, covariance and its properties, spectral decomposition of covariance matrix, whitening. - Continuous RV. Parallels with discrete RVs. Functions of RVs. Mixed RVs. Continuous random vectors. - Appendix: differentiation rules for vectors and matrices. Lec. 4.

Tipo testo	***Testo***
	- Gaussian RVs and their linear transformations. Mahalanobis distance. Classification: - Bayesian prediction: introduction, loss function, conditional risk, argmin/argmax rules - Bayes classification: introduction Lec. 5. Classification - 0/1 loss -> maximum a posteriori (MAP) classifier. Binary MAP. Decision regions. - Classifier performance. - Likelihood ratio tests and receiver operating curve (ROC) - Minimax rule Lec. 6. Classification - Binary Gaussian classification - Homoscedastic case: linear discriminant analysis - Heteroscedastic case: Bhattacharrya bound - Bayes classification with discrete features - Classification with missing data (composite hypothesis testing) Lec. 7. Estimation - Bayesian estimation: introduction - Quadratic loss: minimum mean square error (MMSE) estimator = regression curve - L1 loss: minimum mean absolute error (MMAE) estimator - 0/1 loss: MAP estimator, and maximum likelihood (ML) in uniform prior. - Regression for vector Gaussian case - ML estimation for Gaussian observations Lec. 8. Estimation - ML for multinomial - Conjugate priors in MAP estimation - Estimation accuracy and ML properties, Cramer Rao bounds. Suboptimal (non Bayesian) estimation: - LMMSE estimation (linear regression) - LMMSE derivation with LDU decomposition Lec. 9. Estimation - LMMSE examples - Generalized linear regression - Example: polynomial regression - Sample LMMSE - Generalized sample LMMSE. Lec. 10. Learning - Supervised learning: introduction - Generative vs discriminatie approaches - Example: logistic model - Plug-in learning ML fitting of logistic model: logistic regression Example: handwritten digit recognition. - Bayesian Learning Lec. 11. Learning: - Empirical risk minimization Nonparametric density estimation: - Parzen window estimator - kNN estimator Lec. 12. linear data reduction - Principal component analysis (PCA) - Fisher linear classifier PART 2: Advanced topics and applications (Cagnoni)

Tipo testo	*Testo*
	--- --- Linear Discriminant Analysis: - Support Vector Machines --- Classifier evaluation: - generalization and overfitting (Training/validation/test sets) - performance indices, representations curve, confusion matrices - Classification risk: are all errors equally relevant ? --- Unsupervised classification and clustering - K-means and Isodata algorithms - Self-Organizing Maps - Learning Vector Quantization - Kohonen networks