UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Astrostatistics

2122-1-F5802Q014

Obiettivi

L'uso della statistica e' fra i fondamenti dell'astronomia e dell'astrofisica moderna. Strumenti statistici sempre piu' sofisticati sono ora disponibili su larga scala, spesso indicati coi nomi di "data mining", "machine learning" e "intelligenza artificiale". Il corso fornisce un'introduzione ad alcune di queste tecniche statistiche in una modalita' molto pratica, alternando derivazioni teoriche ed applicazioni computazionali. Nonostante gli esempi trattati verrano presi esclusivamente dal campo dell'astrofisica, questo corso e' appropriato per tutti gli studenti magistrali del corso di laurea in Fisica interessati a tecniche avanzate di statistica e machine learning.

Contenuti sintetici

Programma esteso

Argomenti di statistica e data mining:

- Breve sunto di probabilita' ed inferenza.
- Inferenza Bayesiana (ruolo delle priors, differenze e somiglianze con l'approccio frequentista, selezione di modelli, MCMC).
- Identificazione di strutture nei dati (algoritmi di clustering, estimatori parametrici e non parametrici).
- Riduzione dimensionale (ad esempio Principal Component Analysis).
- Problemi di regressione (overfitting, processi Gaussiani).
- Problemi di classificazione (reti neurali, curve ROC).
- Deep learning.
- Analisi di serie temporali.

Alcuni esempi di dataset astrofisici (lista indicativa):

- Sloan Digital Sky Survey. Dataset con 357 milioni di sorgenti and 1.6 milioni di spettri.
- LINEAR database. Datasets con migliaia di stelle variabili.
- Serie temporali dal Laser Interferometer Gravitational-Wave observatory (LIGO), che ha rivelato piu' di 50 eventi di onde gravitazionali.
- Dati di parallaxes dal satellite GAIA.

Prerequisiti

Non ci sono prerequisiti formali. Una conoscenza di medio livello del linguaggio di programmazione Python è però altamente raccomandata (vedi sotto per alcune risorse di ripasso).

Modalità didattica

Lezioni, 6 crediti.

Materiale didattico

Testo principale:

- "Statistics, Data Mining, and Machine Learning in Astronomy", Željko, Andrew, Jacob, and Gray. Princeton University Press, 2012.
- Software: https://www.astroml.org/

Altre risorse:

- "Statistical Data Analysis", Cowan. Oxford Science Publications, 1997
- "Data Analysis: A Bayesian Tutorial", Sivia and Skilling. Oxford Science Publications, 2006
- "Bayesian Data Analysis", Gelman, Carlin, Stern, Dunson, Vehtari, and Rubin. Chapman & Hall, 2013. Distribuito gratuitamente: http://www.stat.columbia.edu/~gelman/book/
- "Python Data Science Handbook", VanderPlas. O'Reilly Media, 2016. Distribuito gratuitamente: https://jakevdp.github.io/PythonDataScienceHandbook/
- "Practical Statistics for Astronomers", Wall and Jenkins. Cambridge University Press, 2003

- "Bayesian Logical Data Analysis for the Physical Sciences", Gregory. Cambridge University Press, 2005
- "Modern Statistical Methods For Astronomy" Feigelson and Babu. Cambridge University Press, 2012
- "Information theory, inference, and learning algorithms" MacKay. Cambridge University Press, 2003. Distribuito gratuitamente: https://www.inference.org.uk/mackay/itila/book.html
- “Data analysis recipes", Hogg et al.:
  - https://arxiv.org/abs/0807.4820
  - https://arxiv.org/abs/1008.4686
  - https://arxiv.org/abs/1205.4446
  - https://arxiv.org/abs/1710.06068
  - https://arxiv.org/abs/2005.14199

Risorse di ripasso sul linguaggio di programmazione Python:

- https://github.com/jrjohansson/scientific-python-lectures
- https://astrofrog.github.io/py4sci/
- "Learning Scientific Programming with Python", Hill, Cambridge University Press, 2020. https://scipython.com/

Periodo di erogazione dell'insegnamento

Secondo semestre.

Modalità di verifica del profitto e valutazione

L'esame verra' valutato con una prova orale. Un esercizio computazionale verra' assegnato prima dell'esame. Gli studenti dovranno portarlo completato alla prova orale dove verra' discusso col docente. Gli studenti dovranno poi sostenere domande di approfondimento sugli argomenti del corso.

Tutte le lezioni e gli esami si terranno in lingua inglese.

Orario di ricevimento

Sempre, se non mi trovate in ufficio contattatemi via email per un appuntamento.