

Statistica Computazionale e Metodi Multivariati

Schoier Gabriella

Gabriella.schoier@deams.units.it

- **ORARIO**

- Lunedì 16-18 aula tb ed D
- Martedì 14-16 aula tb ed D
- Mercoledì 14-16 aula 4 d ed D

- **Obiettivi**

- Il corso si propone di illustrare i principali aspetti computazionali nell'ambito delle discipline statistiche e del Machine Learning introducendo concetti e strumenti che mettano in grado lo studente di estrarre dalla enorme mole dei dati oggi a disposizione informazioni utili.

2 Aspetti computazionali e gestione / mining dei dati: Data Mining, Web Mining, Spatial Data Mining, Text Mining (cenni)

- Introduzione al Data Mining**
- Dai dati alle informazioni il ruolo dell'Information Technology**
- Il Data Warehouse (cenni)**
- OLAP e OLTP**
- Il processo di Data Mining**

2 Aspetti computazionali e gestione / mining dei dati: Data Mining, Web Mining, Spatial Data Mining, Text Mining (cenni)

- Introduzione al Web Mining :il Web Content Mining , il Web Structure Mining , il Web Usage Mining. Web 2 e Web 3**
- Introduzione al Text Mining**
- Introduzione al Data Mining Spaziale**

3. Analisi Statistica Multivariata Statistical Learning/ Machine Learning e aspetti computazionali:

- L'analisi delle componenti principali**
- L'analisi fattoriale**
- Alcuni metodi di analisi dei cluster**
- Gli Alberi decisionali**
- Le Reti neurali (cenni)**

Wikipedia (cited 5/3/25)

≡ Apprendimento automatico

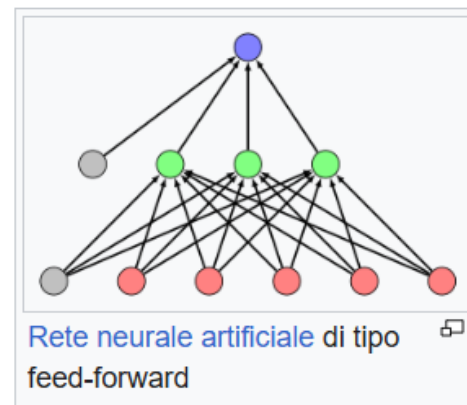
🌐 88 lingue ▾

Voce [Discussione](#)

[Leggi](#) [Modifica](#) [Modifica wikitesto](#) [Cronologia](#) [Strumenti](#) ▾

L'**apprendimento automatico** (in [inglese](#) *machine learning*, abbreviato in **ML**) è una branca dell'[intelligenza artificiale](#) che raccoglie metodi sviluppati negli ultimi decenni del [XX secolo](#) in varie comunità scientifiche, sotto diversi nomi quali: [statistica computazionale](#), [riconoscimento di pattern](#), [reti neurali artificiali](#), filtraggio adattivo, teoria dei [sistemi dinamici](#), [elaborazione delle immagini](#), [data mining](#), [algoritmi adattivi](#), ecc; che utilizza metodi [statistici](#) per migliorare la performance di un [algoritmo](#) nell'identificare pattern nei dati.

Nell'ambito dell'[informatica](#), l'apprendimento automatico è una variante alla [programmazione](#) tradizionale nella quale in una macchina si predispone l'abilità di apprendere qualcosa dai dati in maniera autonoma, senza istruzioni esplicite^{[1][2]}. Gli algoritmi di apprendimento automatico sono utilizzati in un'ampia varietà di branche del sapere, come la [medicina](#), il filtraggio delle [e-mail](#), il [riconoscimento vocale](#) e la [visione artificiale](#), dove è difficile o non fattibile sviluppare algoritmi convenzionali per eseguire i compiti richiesti^[3].



Wikipedia (cited 5/3/25)

Un'altra categorizzazione dei compiti dell'apprendimento automatico si rileva quando si considera l'output desiderato del sistema di apprendimento automatico.^[22]

- Nella **classificazione**, gli output sono divisi in due o più classi e il sistema di apprendimento deve produrre un modello che assegni gli input non ancora visti a una o più di queste. Questo viene affrontato solitamente in maniera supervisionata. Il filtraggio anti-spam è un esempio di classificazione, dove gli input sono le email e le classi sono "spam" e "non spam".
- Nella **regressione**, che è anch'essa un problema supervisionato, l'output e il modello utilizzati sono **continui**. Un esempio di regressione è la determinazione della quantità di olio presente in un oleodotto, avendo le misurazioni dell'attenuazione dei raggi gamma che passano attraverso il condotto. Un altro esempio è la predizione del valore del tasso di cambio di una valuta nel futuro, dati i suoi valori in tempi recenti.^[23]
- Nel **clustering** un insieme di input viene diviso in gruppi. Diversamente da quanto accade per la classificazione, i gruppi non sono noti prima, rendendolo tipicamente un compito non supervisionato. Un tipico esempio di clustering è l'analisi del comportamento degli utenti di un sito web.

- **4. Aspetti computazionali nel caso di Metodi Statistici lineari**
- **5. Aspetti computazionali nel caso di Metodi Statistici non lineari**

- **6. Simulazione e metodi di ricampionamento (cenni)**
 - **Il Metodo Monte Carlo e alcune sue applicazioni**
 - **Introduzione ai metodi di ricampionamento: il metodo Bootstrap e il metodo Jackknife**

- **Formano parte integrante applicazioni con l'utilizzo del linguaggio SAS (SAS Studio)**

- **Verifica dell'apprendimento**

L'esame consiste di una prova scritta/pratica con domande la cui risposta prevede l'utilizzo del linguaggio SAS e una prova orale su tutto il programma del corso.

Linguaggio SAS

- Versione sas disponibile su macchina virtuale
- Versione sas studio disponibile on line previa registrazione



Learn

Learn SAS programming from basic to advanced techniques using our free cloud-based learning environment.

- Available anytime with minimal downtime to use at your convenience.
- Free access to SAS Studio, which supports all major desktop operating systems.
- 5GB of data storage for your personal files.
- Compatible with free [e-Learning courses](#) for guided instruction.
- Ideal for certification training.
- Access to free online [SAS Communities](#), which provides access to more learning resources and a community of fellow learners sharing knowledge and advice.



Teach

Additional resources also available for use in qualified classrooms. See the [license](#) and the [terms of use and conditions](#) for more information on qualifications and restrictions.

- Additional 3GB of data storage per instructor to share files with students.
- Additional software options available when using Microsoft Windows:
 - Learn SAS programming with SAS® Enterprise® Guide.
 - Teach data mining techniques using SAS® Enterprise® Miner.
 - Teach forecasting using SAS® Forecast® Studio®.
- Join the [SAS Educator Portal](#) for 24/7 access to free instructional materials, e-learning and certification exam prep to make it easier than ever to teach analytics.

Linguaggio SAS: registration instruction

- [Vedere su Moodle2](#)  registration-sas-on-demand-for-academ...
- <https://welcome.oda.sas.com/>
- [https://www.sas.com/it it/software/on-demand-for-academics/features-list.html](https://www.sas.com/it_it/software/on-demand-for-academics/features-list.html)



Learn

Learn SAS programming from basic to advanced techniques using our free cloud-based learning environment.

- Available anytime with minimal downtime to use at your convenience.
- Free access to SAS software and training materials.
- 5GB of data storage for your learning projects.
- Compatible with various operating systems.
- Ideal for certification preparation.
- Access to free resources and advice.



Teach

Additional resources and terms of use and conditions.

- Additional 30-day trial period.
- Additional software licenses.
- Learn SAS programming from basic to advanced techniques using our free cloud-based learning environment.
- Teach data mining techniques using SAS® Enterprise® Miner.
- Teach forecasting using SAS® Forecast® Studio®.

Sign In



☐ Accept the terms of the [license](#) and the [terms of use and conditions](#).

Sign In

[Forgot Password?](#)

[Don't have a SAS Profile?](#)



Linguaggio SAS

- https://www.sas.com/it_it/software/on-demand-for-academics/features-list.html

SAS® OnDemand for Academics: Studio

Available to professors, students and independent learners.

- Access the power of SAS Analytics simply by connecting to a website.
- Use the pre-defined tasks to generate SAS code.
- Submit code from a wide range of devices, from wherever they happen to be.
- Avoid the need to download and install a software application or updates.
- Teach or enhance SAS skills using traditional SAS programming windows (Program Editor, Log, Output).
- Create new or upload existing SAS program files.
- Access SAS software, including:
 - SAS/STAT® – State-of-the-art statistical software.
 - Base SAS® – Flexible, extensible fourth-generation programming language for data access, transformation and reporting.
 - SAS/IML® – Interactive matrix programming and exploratory data analysis.
 - SAS/OR® – Enhanced operations research methods for optimizing processes and addressing management science challenges.
 - SAS/QC® – Tools for statistical quality improvement.
 - SAS/ETS® – Model, forecast and simulate processes with econometric and time series analysis.

Riferimenti bibliografici

Hastie T., Tibshirani R. J. ,Friedman J. (2025), Gli elementi dell'apprendimento: Data Mining , Inferenza e Previsione, a cura di Biganzoli E., Boracchi P., Gaito S., Nicolussi F., Salini S., Ambrogi F., Manzi G., Marano G.,Zignani M., Piccin Editore.

Gareth J., Witten D., Hastie T., Tibshirani R. J. (2021), Introduzione all'apprendimento statistico, a cura di Salini S., Gaito S., Boracchi P., Ambrogi F., Manzi G., Biganzoli E., Piccin Editore.

Hastie T., Tibshirani R. J., Friedman J., (2017), The Elements of Statistical Learning Data Mining, Inference, and Prediction. Second edition Springer

Zani S., Cerioli A., (2007), Analisi dei dati e data mining per le decisioni aziendali, Giuffrè Editore.

Giudici P.,(2005), Data Mining metodi informatici , statistici e applicazioni. Seconda edizione McGraw-Hill

Roiger R. J., Geatz M. W., (2003),Introduzione al Data Mining, McGraw-Hill

Thisted R. A., (2000), Elements of Statistical Computing,