



WORKSHOP

MACHINE LEARNING IN STATA: UN'INTRODUZIONE

Roma, 20-22 Novembre 2019

Gli ultimi anni hanno visto una disponibilità senza precedenti di informazioni su fenomeni sociali, economici e relativi alla salute. Ricercatori, professionisti e responsabili delle politiche hanno oggi accesso a enormi dataset (i cosiddetti *Big Data*) su persone, aziende e istituzioni, dispositivi cellulari, web, satelliti, ecc., con sempre maggiore dettaglio.

Il *machine learning* (o “apprendimento automatico”) è un approccio relativamente nuovo all’analisi dei dati, che si colloca nell’intersezione tra statistica, informatica e intelligenza artificiale. Il suo obiettivo principale è quello di trasformare le informazioni in conoscenza e valore, “lasciando che i dati parlino da soli”.

A tal fine, il *machine learning* limita le ipotesi preliminari sulla struttura dei dati e fa affidamento su una filosofia che supporta lo sviluppo di algoritmi, di procedure computazionali e d’ispezione grafica dei risultati più che su assunzioni analitiche e soluzioni algebriche. Computazionalmente irrealizzabile fino a pochi anni fa, il *machine learning* è un prodotto dell’era del computer, della potenza di calcolo e della capacità di apprendimento delle macchine odierne, dello sviluppo dell’hardware e del continuo sviluppo del software.

Questo corso è una introduzione alle tecniche di *machine learning* che utilizzano Stata. Stata possiede oggi vari pacchetti per eseguire il *machine learning* che sono tuttavia poco conosciuti da molti suoi utenti. Questo corso colma questa lacuna rendendo i partecipanti più familiari (e meglio informati) del potenziale di Stata per trarre conoscenza e valore dai dati, possibilmente di grandi dimensioni e “rumorosi”. L’approccio all’insegnamento sarà principalmente basato sul linguaggio grafico e sull’intuizione più che sull’algebra. Le lezioni si avvarranno di esempi sia simulati che reali, e permetterà di bilanciare equamente sessioni teoriche e sessioni pratiche.

Dopo il corso, i partecipanti avranno una migliore comprensione del potenziale di Stata per eseguire il *machine learning*, diventando così in grado di padroneggiare compiti di ricerca che includono, tra gli altri: (i) rilevamento d’importanza dei fattori, (ii) estrazione segnale-rumore, (iii) corretta specificazione del modello, (iv) classificazione senza modello, sia da un punto di vista di data mining che di approccio causale.

In linea con la nostra filosofia di formazione, le sessioni teoriche sono affiancate da illustrazioni pratiche ed esempi provenienti da diverse discipline sia sociali che biomediche, in cui il docente chiarisce le limitazioni e i punti di forza di ogni metodologia, nonché i criteri per la scelta e l’implementazione dello strumento di analisi statistica più appropriato per il problema oggetto di studio.

CODICE WORKSHOP

I-EF34

DATA E LUOGO

Roma, 20-22 Novembre 2019

DESTINATARI

Il corso è di interesse per ricercatori e analisti in economia, medicina, marketing e scienze sociali che desiderano acquisire gli strumenti fondamentali per implementare l’approccio di *machine learning* sui così detti *Big Data*.

MACHINE LEARNING IN STATA: UN'INTRODUZIONE

SESSIONE IV: SELEZIONE DEL MODELLO E REGOLARIZZAZIONE

1. Metodi basati sulla "subset selection"
 - Selezione di sottoinsiemi ottimali con **combinatorics**
2. Metodi di shrinkage
 - Lasso, Ridge, and Elastic regression con **lassopack**
3. Sensitivity analysis
 - Incertezza del modello e sua credibilità
 - L'algoritmo LOCO per l'analisi di sensitività con **sensimatch**

SESSIONE V: REGRESSIONE E CLASSIFICAZIONE "AD ALBERO"

1. Un'introduzione agli alberi di regressione e classificazione
2. *Bagging*, foreste casuali e **boosting**
3. Il comando Stata **stree**

QUOTA DI ISCRIZIONE

La partecipazione al workshop è soggetta al pagamento della seguente quota di iscrizione:

Studente*: € 790.00

Assegnista / Specializzando: € 1050.00

Università: € 1310.00

Commerciale: € 1645.00

*Per usufruire dello status "studente" è necessario presentare copia del libretto universitario o un certificato di iscrizione (in carta semplice) all'Università ed essere *studenti a tempo pieno*. Studenti lavoratori dovranno considerare la tariffa Assegnisti / Specializzandi.

I prezzi si intendono IVA 22% esclusa. L'aliquota IVA non sarà applicata per Enti Pubblici soggetti ad esenzione a norma dell'art. 14 c. 10 della L. 537/93 per la partecipazione a corsi di formazione dei propri dipendenti.

La quota di iscrizione include il pranzo, il materiale didattico e una licenza temporanea del software Stata (si consiglia di venire muniti del proprio computer o di chiedere informazioni alla segreteria per l'eventuale noleggio, al momento dell'iscrizione). Dà inoltre diritto ad uno sconto sull'acquisto di una nuova licenza per singolo utente del Software Stata (ad esclusione della versione per Studenti) e sull'acquisto di testi in catalogo Stata Press.

L'iscrizione al corso dovrà avvenire tramite lo specifico modulo di registrazione e pervenire a TStat S.r.l. entro il 2 Novembre 2019. Lo svolgimento è condizionato dal raggiungimento di un numero minimo di 8 partecipanti ed un numero massimo di 15.

Ulteriori informazioni sulla modalità di iscrizione, incluso termini e condizioni di partecipazione sono disponibili nel nostro sito alla pagina <https://www.tstat.it/formazione/machine-learning-stata/>

TESTI UTILI

Econometric Analysis of Cross Section and Panel Data, Wooldridge, MIT Press (2010)

Microeconometrics Using Stata, Cameron e Trivedi, Revised Edition, StataPress (2010)

The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Hastie, T., Tibshirani, R., Friedman, J., Springer (2009)

An Introduction to Statistical Learning, Gareth, J., Witten, D., Hastie, T., Tibshirani, R., Springer (2013)

CONTATTI

Monica Gianni

TStat S.r.l. | Via Rettangolo, 12-14
I-67039 Sulmona (AQ)
T. +39 0864 210101

TStat Training | Kleebergstraße, 8
D-60322 Frankfurt am Main

formazione@tstat.it

www.tstat.it

www.tstattraining.eu

The logo for TStat, featuring a large, stylized red 'T' followed by the word 'Stat' in a black, cursive script font.

www.tstattraining.eu | www.tstat.it | www.tstat.eu