

## AI per l'analisi e previsione delle serie temporali

CODICE	DT0275
DURATA	2 gg
PREZZO	1.350,00 €
EXAM	

### DESCRIZIONE

---

Ogni azienda gestisce nel suo business delle serie temporali, cioè delle serie di misurazioni nel tempo di un particolare fenomeno. L'azienda ha in genere la necessità di capire la serie e soprattutto di prevederne i prossimi valori. Purtroppo le serie temporali sembrano spesso ingannevolmente semplici.

L'analisi/previsione delle serie temporali è un insieme di tecniche statistiche, applicabili ad una grande varietà di situazioni di business, per indagare una serie storica temporale (*time series analysis*), determinarne il processo che è alla base e trarne previsioni per il futuro, più o meno prossimo (*time series forecasting*). L'analisi capisce e spiega il passato, il forecasting prevede il futuro. Il processo ha una parte deterministica, composta di componenti tendenziali, cicliche e/o stagionali, ed una parte casuale che è descrivibile mediante un modello probabilistico.

Il corso illustra le nuove tecniche di Artificial Intelligence e Machine Learning che negli ultimissimi tempi sono emerse e si sono affermate e che permettono una più efficace previsione delle serie temporali.

L'analisi/previsione delle serie temporali è frequentemente usata nel *Sales & Marketing* (ad es. per prevedere vendite e prezzi dei prodotti), nel *Finance* (ad es. per prevedere i prezzi azionari di chiusura), nel *Transportation* (ad es. per prevedere il numero di viaggiatori nel futuro e la relativa domanda), e nel mercato immobiliare (per prevedere i prezzi futuri delle case), e quindi i casi d'uso del corso saranno relativi soprattutto a questi mercati. Le tecniche illustrate in questo corso, comunque, sono generali ed applicabili a qualsiasi mercato verticale, non solo i quattro indicati.

Il corso affronta il tema in modo completamente *hands-on*, tramite molti esempi e casi d'uso direttamente col codice, tutto commentato in italiano, senza l'ausilio di slide.

### OBIETTIVI RAGGIUNTI

---

Al termine del corso i partecipanti saranno in grado di:

- avere una solida base teorica e pratica sull'analisi e previsione delle serie temporali
- applicare le tecniche di analisi e previsione delle serie temporali ai problemi reali delle aziende
- conoscere i diversi casi d'uso delle serie temporali nei vari settori aziendali

- utilizzare il codice software fornito nel corso come punto di partenza per adattarlo alla propria realtà aziendale.

## TARGET

---

Il corso si rivolge a Data Scientist, Analisti di rischio, Project Manager, Sviluppatori e chiunque voglia acquisire una competenza effettiva e pratica sul tema del rilevamento delle frodi tramite le nuove tecniche di AI.

## PREREQUISTI

---

Conoscenze di base effettiva del linguaggio Python od R

## CONTENUTI

---

### Introduzione

- Stabilire obiettivo e contesto dell'analisi temporale: descrizione o previsione? l'orizzonte temporale? la scala temporale?
- Previsione vs. forecast;
- Pre-processing dei dati grezzi: anomalie, discontinuità, differenti intervalli, differenti metriche, ecc;
- Esplorare e visualizzare una serie temporale;
- Lo strumento principale: il time plot della serie;
- La decomposizione in trend, cicli e stagionalità di una serie temporale;
- La differenza tra stagione e "stagionalità";
- L'importanza grafica dei colori per meglio comprendere la serie;
- Il "rumore" dalla serie: come analizzarlo;
- Strumenti utili: zoom, cambio di scala, linee di trend, soppressione stagionalità;
- Esame e trattamento degli outlier;
- Esame del correlogramma;
- Tecniche di visualizzazione avanzata della serie.
- Il processo di forecasting: gli step;
- Previsione *data-driven* anziché *model-driven*: criteri di scelta.

### La valutazione predittiva (sul futuro) di una serie temporale

- Il partizionamento dei dati (in training, validation ed eventualmente test);
- Stimare gli errori delle previsioni con le metriche MAE, MAPE e RMSE;
- L'importanza della visualizzazione per valutare e confrontare modelli differenti;
- Confrontare modelli di forecasting differenti e valutare in modo comparato le loro capacità predittive;
- Un benchmark di riferimento per qualsiasi serie: Naive Forecast e Seasonal Naive Forecast;
- Ricombinare i dataset per prevedere meglio;
- Aumentare l'orizzonte della previsione (*forecast horizon*).

## Prevedere i prossimi valori della una *time series* con la regressione lineare multivariata

- individuare i predittori più adatti per catturare trend e/o stagionalità della serie;
- fittare il modello di regressione ai dati disponibili;
- prevedere i prossimi valori della serie tramite l'equazione della regressione;
- i tipi di trend più frequenti: lineare, esponenziale e polinomiale;
- i due tipi di stagionalità: additiva e moltiplicativa;
- prevedere con trend e stagionalità insieme nel modello;
- quantificare l'auto-correlazione tra i valori della serie (oltre il trend e la stagionalità);
- costruire un modello auto-regressivo con la tecnica della ARIMA per previsioni più precise;
- valutare se la serie è un semplice random walk (tramite l'ARIMA);
- valutare la capacità predittiva del modello sul futuro (tramite i dati di validazione e la visualizzazione).

## Prevedere i prossimi valori della una *time series* con il livellamento (*smoothing*)

- livellatori per serie senza trend e stagionalità: moving average e simple exponential smoother;
- impostare la lunghezza della serie ed i suoi pesi;
- visualizzazione di una serie con il moving average;
- livellatori per serie con trend e/o stagionalità:
- rimozione di trend/stagionalità;
- *advanced exponential smoother*: double, Holt-Winter.
- impostare le costanti di livellamento (smoothing constants) per determinare la velocità di adattamento ai nuovi dati: scelte e significati;
- valutare la capacità predittiva della serie sul futuro (tramite i dati di validazione).

## Altre tecniche

- Combinazione di metodi per previsioni più robuste e precise: two-level forecasting, ensemble;
- Clustering di una serie temporale;
- Classificazione di una serie temporale etichettata con 'knn' ed altre tecniche;
- Reti neurali?

## Esplorare le TS in modo interattivo

Analizzare l'impatto di un evento sulla *time series*: come si fa

Time series multiple: cenni.

Time series utilizzate nel corso:

- vendite di prodotti, viaggiatori in aereo ed in treno, domanda di traffico, prezzi azionari di chiusura, spedizioni di merci, ricavi, prezzi delle case, ore lavorate dal personale.