



UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Dynamic Models and Machine Learning Forecasting M

2627-1-F8206B035

Obiettivi formativi

L'obiettivo dell'insegnamento è quello di fornire le basi teoriche e applicative dei modelli lineari dinamici per serie storiche e di *machine learning forecasting*. Si illustreranno i fondamenti della teoria della previsione, la forma state space e i metodi di filtraggio collegati, oltre alle tecniche di machine learning utili per prevedere serie storiche complesse.

Il corso viene impartito in modalità *blended learning* con videolezioni e lezioni frontali in laboratorio informatico. Tutti i modelli vengono implementati su dati e problemi economici reali usando l'ambiente open source R.

L'insegnamento completa la preparazione statistico-economica degli studenti di tutti i percorsi del corso di laurea magistrale, fornendo strumenti statistici per lavorare con serie storiche macroeconomiche, aziendali, finanziarie e di altra natura.

Al termine dell'insegnamento, i risultati di apprendimento attesi sono i seguenti:

Conoscenza e capacità di comprensione (DdD 1): Lo studente sarà in grado di costruire modelli lineari dinamici ad-hoc per prevedere e filtrare serie storiche reali. Inoltre, sarà in grado di utilizzare metodi di machine learning a fini di previsione di serie storiche e di produrre le *features* più adatte affinché tali modelli apprendano alcune caratteristiche delle serie storiche quali trend e stagionalità.

Conoscenza e capacità di comprensione applicate (DdD 2): Lo studente sarà in grado di utilizzare in modo avanzato il linguaggio R e i pacchetti relativi alla forma state space e ad alcuni modelli di machine learning per implementare soluzioni a problemi reali.

Autonomia di giudizio (DdD 3): Lo studente saprà analizzare criticamente i risultati dei modelli, valutandone l'accuratezza e la coerenza rispetto agli obiettivi.

Abilità comunicative (DdD 4): Lo studente acquisirà la capacità di comunicare i risultati delle analisi a esperti del settore utilizzando un linguaggio tecnico appropriato, sia in italiano sia in inglese

Capacità di apprendere (DdD 5): Lo studente svilupperà le competenze necessarie per approfondire autonomamente nuovi modelli e procedure computazionali indispensabili per l'attività professionale di statistico e *forecaster*.

Contenuti sintetici

- Teoria della previsione
- Forma state space dei modelli lineari dinamici e filtro di Kalman
- Modelli UCM
- Machine learning per forecasting
- Applicazioni a dati reali con R

Programma esteso

- Previsore ottimo
- Previsore lineare ottimo
- Forma state space / modelli lineari dinamici
- Principali componenti dei modelli UCM (trend, ciclo, stagionalità)
- Modelli ARIMA e UCM in forma state space
- Filtro di Kalman e stima di massima verosimiglianza
- Inizializzazione delle variabili di stato
- Smoothing delle variabili di stato e dei disturbi
- Principali algoritmi di machine learning (ML)
- Problemi di estrapolazione di molti algoritmi di ML e soluzioni
- La costruzione delle *features* temporali per i modelli di ML
- Esercizi e casi di studio usando R

Prerequisiti

Buone conoscenze di inferenza statistica, algebra matriciale e fondamenti di serie storiche (processi stazionari, processi integrati, modelli ARIMA).

Conoscenze di base di R.

Metodi didattici

Il corso viene impartito in modalità *blended learning*: il 50% dell'insegnamento avviene in presenza (in laboratorio) e il 50% avviene in remoto per mezzo di video-lezioni, applicazioni web, test ed esercizi on-line e forum di domande e risposte.

Le lezioni in presenza sono sempre in laboratorio e rendono operativa, per mezzo di applicazioni a dati reali, la teoria acquisita autonomamente dallo studente per mezzo delle videolezioni, del manuale e degli altri ausili didattici presenti nella pagina elearning. Inoltre, le lezioni in presenza sono l'occasione per gli studenti di esporre i propri

dubbi e proporre le proprie domande su quanto appreso autonomamente, in modo che il docente possa fornire spiegazioni alternative per quanto non sia risultato chiaro.

La durata delle lezioni in presenza è di due o tre ore per un totale complessivo di 23 ore.

Le videolezioni sono di didattica erogativa, e diversi questionari di autovalutazione e web-app permettono di verificare e migliorare l'apprendimento. Le lezioni in presenza sono un mix di didattica erogativa (circa il 30%) e didattica interattiva (circa il 70%).

L'insegnamento è erogato in lingua inglese.

Modalità di verifica dell'apprendimento

L'esame si compone di una prova scritta su domande teoriche ed esercizi (durata 1 ora) e di una prova pratica in laboratorio usando R (durata 1 ora). Il risultato di ciascuna prova concorre al 50% del voto finale.

Dettaglio delle prove:

- **Prova scritta:** verifica le conoscenze teoriche relative alla teoria della previsione, ai modelli a componenti non osservabili, alla forma state space e al machine learning forecasting attraverso cinque domande aperte. Le risposte concorrono con il medesimo peso alla valutazione dello scritto.
- **Prova pratica:** richiede la costruzione e stima di modelli su una serie storica fornita, inclusa la produzione delle diagnostiche richieste.
- **Progetto di gruppo:** il docente può proporre agli studenti frequentanti di sostituire la prova pratica con un progetto originale di gruppo.

Competenze valutate:

Capacità di comprensione teorica della modellistica state space e dell'uso dei modelli di machine learning per il forecasting (DdD 1).

Abilità nel risolvere problemi di previsione ed estrazione di segnali tramite codice R (DdD 2).

Capacità di analisi critica e diagnostica dei risultati (DdD 3).

Criteri di valutazione e graduazione del voto: La valutazione tiene conto della precisione delle risposte teoriche (50%) e della corretta implementazione e interpretazione dei modelli statistici nella prova pratica (50%). L'eccellenza (30 e lode) richiede il pieno raggiungimento di tutti gli obiettivi formativi e una particolare padronanza nell'uso degli strumenti computazionali.

Prove in itinere: Non sono previste prove parziali obbligatorie durante il corso, ma agli studenti frequentanti possono essere proposti esercizi facoltativi da fare a casa che possono contribuire al voto finale.

Testi di riferimento

Pelagatti (2015) *Time Series Modelling with Unobserved Components*, Chapman and Hall/CRC (scaricabile gratuitamente sotto indirizzo IP di Bicocca)

Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. [OTexts.com/fpp2](https://otexts.com/fpp2)

Appunti del docente per la parte di machine learning forecasting

Periodo di erogazione dell'insegnamento

IV ciclo (maggio - giugno)

Lingua di insegnamento

Italiano

Sustainable Development Goals
