

Analisi Dati per la Business Intelligence

II edizione 2013/2014



UNIVERSITÀ DEGLI STUDI
DI TORINO

**Dipartimento di Culture,
Politica e Società
Dipartimento di Informatica**

gestito da



In collaborazione con



PROGRAMMA DIDATTICO

1. Software per la Business Intelligence

36 ore di lezione frontale, 86 ore di studio individuale

Crediti Formativi: 5

I temi trattati nel presente insegnamento saranno:

<ul style="list-style-type: none">• Business Intelligence: individuazione, archiviazione, analisi e comunicazione dei dati, Big Data
<ul style="list-style-type: none">• Backend: Strumenti per archiviazione e aggiornamento dei dati, flat tables, DW e DataMart
<ul style="list-style-type: none">• Il percorso analitico del data Mining: Semma in SAS
<ul style="list-style-type: none">• I software proprietari: BO, SAS e webreporting
<ul style="list-style-type: none">• Gli open source based software: Rapid Miner, Knime
<ul style="list-style-type: none">• Graphics: Analytics, Communication tools, Dashboards, Motion
<ul style="list-style-type: none">• Modelli di analisi testuale: Text Mining
<ul style="list-style-type: none">• Sentiment analysis (analisi dei blog e dei social network)

2. DB Management

24 ore di lezione frontale, 51 ore di studio individuale

Crediti Formativi: 3

I temi trattati nel presente insegnamento saranno:

<ul style="list-style-type: none">• Costruire le basi informative a partire dai bisogni di conoscenza del cliente
<ul style="list-style-type: none">• Esplorare le fonti informative interne ed esterne all'azienda
<ul style="list-style-type: none">• Capacità di ampliare la visione del cliente (dall'applicazione di BI ad un'architettura enterprise di Data Warehouse)
<ul style="list-style-type: none">• Come costruire basi dati decisionali (modellare i dati decisionali, la costruzione della base dati decisionale)
<ul style="list-style-type: none">• "Lessons learned" nella costruzione di basi dati decisionali

3. Raccolta dei dati e data quality

36 ore di lezione frontale, 89 ore di studio individuale

Crediti Formativi: 5

I temi trattati nel presente insegnamento saranno:

<ul style="list-style-type: none">• Prestare attenzione alla qualità delle informazioni (le dimensioni di qualità dei dati)
<ul style="list-style-type: none">• La gestione del data anagrafico per il riconoscimento e la contattabilità del Cliente
<ul style="list-style-type: none">• L'assessment di qualità di una base dati
<ul style="list-style-type: none">• Costruire processi di raccolta e bonifica dati
<ul style="list-style-type: none">• Progettare i flussi di aggiornamento e snellimento delle basi dati decisionali
<ul style="list-style-type: none">• "Lessons learned" sulla raccolta e qualità dei dati

4. Statistica per la BI

24 ore di lezione frontale, 51 ore di studio individuale

Crediti Formativi: 3

I temi trattati nel presente insegnamento saranno:

<ul style="list-style-type: none">• Variabili e misure
<ul style="list-style-type: none">• Elementi di base per lo studio dell'inferenza statistica: calcolo combinatorio, probabilità
<ul style="list-style-type: none">• Campionamento
<ul style="list-style-type: none">• Le distribuzioni statistiche

5. Mining I – Statistica Descrittiva

24 ore di lezione frontale, 51 ore di studio individuale

Crediti Formativi: 3

I temi trattati nel presente insegnamento saranno:

<ul style="list-style-type: none">• I dati, le funzioni, i grafici, altri oggetti
<ul style="list-style-type: none">• Formati e tipologie di dati; la rappresentazione grafica
<ul style="list-style-type: none">• Indici statistici e loro applicazioni
<ul style="list-style-type: none">• Tabelle di contingenza e indipendenza statistica
<ul style="list-style-type: none">• La correlazione e la regressione

6. Mining II – Relazioni Statistiche

24 ore di lezione frontale, 51 ore di studio individuale

Crediti Formativi: 3

I temi trattati nel presente insegnamento saranno:

<ul style="list-style-type: none">• Introduzione al processo di estrazione di conoscenza dai dati (KDD)
<ul style="list-style-type: none">• La classificazione e metodi predittivi
<ul style="list-style-type: none">• Sperimentazione su metodi predittivi con software open source
<ul style="list-style-type: none">• Apprendimento non supervisionato e clustering
<ul style="list-style-type: none">• Sperimentazione su metodi non supervisionati con software open source

7. Mining III – Sintesi dell'informazione statistica

24 ore di lezione frontale, 51 ore di studio individuale

Crediti Formativi: 3

I temi trattati nel presente insegnamento saranno:

<ul style="list-style-type: none">• Conoscere i propri clienti costruendone il profilo attraverso la Segmentazione
<ul style="list-style-type: none">• Comportamentale: identificazione dell'obiettivo e impostazione dell'analisi
<ul style="list-style-type: none">• Sintesi dei fenomeni multidimensionali: analisi fattoriale e analisi delle corrispondenze
<ul style="list-style-type: none">• Individuazione di differenti tipologie di clienti: analisi cluster
<ul style="list-style-type: none">• Interpretazione dei risultati, passaggio in produzione e supporto alle strategie di marketing dell'azienda
<ul style="list-style-type: none">• Esempi di sintesi dell'informazione statistica in ambito Pubblica Amministrazione (indicatori e cruscotti di performance, data e text mining, visualizzazione di informazioni)

8. Modelli e modelli di equazioni strutturali

24 ore di lezione frontale, 51 ore di studio individuale

Crediti Formativi: 3

I temi trattati nel presente insegnamento saranno:

<ul style="list-style-type: none">• Modelli di regressione lineare e regressione logistica
<ul style="list-style-type: none">• Identificazione di relazioni spurie
<ul style="list-style-type: none">• Modelli PATH
<ul style="list-style-type: none">• Modelli a variabili latenti - Applicazioni in ambito Marketing e CRM

9. Tecniche di simulazione

24 ore di lezione frontale, 51 ore di studio individuale
Crediti Formativi: 3

I temi trattati nel presente insegnamento saranno:

<ul style="list-style-type: none">• Scelta dei modelli (lineari, non lineari)
<ul style="list-style-type: none">• Stima dei modelli
<ul style="list-style-type: none">• Tuning dei modelli
<ul style="list-style-type: none">• Previsioni
<ul style="list-style-type: none">• Applicazioni in campo HR e Marketing

10. Interpretazione e comunicazione delle relazioni statistiche

60 ore di lezione frontale, 140 ore di studio individuale

Crediti Formativi: 8

I temi trattati nel presente insegnamento saranno:

- Media, Pubblici e stakeholder della ricerca scientifica: la loro rilevanza nell'impostazione teorica e nella comunicazione dei risultati e nella comunicazione del metodo (probabilismo, falsificazionismo)

- Modi efficaci di comunicazione dei risultati nell'ambito Business Intelligence

- Responsabilità giuristica della ricerca, proprietà dei dati del finanziatore, copy right e open access nella sua diffusione

- Informazioni per la Business Intelligence in impresa

- I numeri nella comunicazione

- Numeracy: valutare i numeri e ciò che ci viene detto con i numeri

- Banco di prova (benchmarking): costruire riferimenti empirici per la valutazione dei numeri

- IntenSionality: il contesto in cui i numeri vengono prodotti e i modelli "teorici" da cui derivano

- Presentazione a diverse audience : Slide, Video, ecc.