

Machine Learning and Data Mining

Giorgio Medico
Carlo Gervasi

fall semester 2024

Contents

Chapter 1

Data Mining

1.1 Introduzione alla Business Intelligence

La Business Intelligence rappresenta un insieme completo di metodologie, processi e tecnologie finalizzate alla trasformazione dei dati aziendali in informazioni utili per il processo decisionale.

1.1.1 Definizioni Fondamentali

Definizione Gartner: “La Business Intelligence è un termine ombrello che include applicazioni, infrastrutture, strumenti e best practice che abilitano l’accesso e l’analisi delle informazioni per migliorare decisioni e performance.”

Definizione Forrester: “La Business Intelligence è un insieme di metodologie, processi, architetture e tecnologie che trasformano i dati grezzi in informazioni significative e utili.”

1.1.2 Obiettivi Principali

Gli obiettivi fondamentali della BI sono:

- Trasformare i dati grezzi in informazioni utili
- Supportare strategie aziendali consapevoli
- Fornire le informazioni giuste alle persone giuste
- Garantire la tempestività dell’informazione
- Ottimizzare i canali di distribuzione delle informazioni

1.2 Il Data Warehouse

1.2.1 Caratteristiche Fondamentali

Un Data Warehouse è una collezione di dati che supporta i processi decisionali e presenta le seguenti caratteristiche distintive:

Orientamento ai soggetti Focus su concetti specifici dell’impresa come clienti, prodotti, vendite

Integrazione Unificazione di dati provenienti da fonti eterogenee

Non volatilità I dati, una volta inseriti, non vengono modificati

Variazione temporale Mantenimento della storicità dei dati

1.2.2 Vantaggi dei Sistemi DWH

I principali vantaggi offerti da un sistema DWH sono:

- Gestione efficiente dei dati storici
- Esecuzione di analisi multidimensionali
- Semplicità di apprendimento per gli utenti
- Supporto al calcolo di indicatori prestazionali

1.3 La Piramide della BI

La struttura della Business Intelligence può essere rappresentata come una piramide con quattro livelli:

Dati Database operazionali e fonti dati

Informazione OLAP e data warehouse

Conoscenza Data Mining e modelli di apprendimento

Decisioni Analisi what-if e modelli di simulazione

1.4 Campi di Applicazione

I sistemi DWH trovano applicazione in numerosi settori:

1.4.1 Commercio

- Analisi delle vendite e dei reclami
- Controllo spedizioni e inventario
- Customer care e relazioni pubbliche

1.4.2 Servizi Finanziari

- Analisi del rischio
- Gestione carte di credito
- Rilevamento frodi

1.4.3 Altri Settori

- **Trasporti:** Gestione veicoli
- **Telecomunicazioni:** Analisi profilo clienti e performance di rete
- **Sanità:** Analisi ammissioni e dimissioni pazienti

1.5 Infrastruttura BI

Un'efficace piattaforma BI richiede:

- Hardware dedicato
- Infrastruttura di rete
- Database ottimizzati
- Data Warehouse
- Software di front-end per la visualizzazione

Chapter 2

Machine Learning