

[K1.9] DATA SCIENTIST**Descrizione sintetica:**

Il Data scientist si occupa di modellizzare problemi complessi, a supporto della presa di decisioni, progettando modalità di analisi matematica e statistica, relativi a dati a differente strutturazione, fra cui i big data. Individua, sviluppa ed applica metodi e tecniche di raccolta, organizzazione e standardizzazione di dati, sui quali applica protocolli di data mining, rivolti alla produzione di analytics descrittivi, diagnostici, predittivi e prescrittivi. Implementa algoritmi e software di analisi e machine learning applicato. Comunica i risultati delle analisi, individuando ed applicando tecniche di data visualization. Definisce ed implementa le caratteristiche organizzative dei processi stabili di data analytic.

SISTEMI DI REFERENZIAMENTO	
Sistema di riferimento	Denominazione
Settore economico-professionale (S.E.P.)	Servizi digitali
Area/e di Attività (AdA) del Repertorio nazionale delle qualificazioni regionali a cui il profilo afferisce	ADA.14.01.20 - Data Science and Analytics
Livello E.q.f.	6
Posizione classificatoria ISTAT CP 2011	2.1.1.5.2 - Analisti e progettisti di basi dati
Posizione/i classificatoria/e ISTAT ATECO 2007	62.01.00 - Produzione di software non connesso all'edizione 62.02.00 - Consulenza nel settore delle tecnologie dell'informatica 62.09.09 - Altre attività dei servizi connessi alle tecnologie dell'informatica nca 63.11.20 - Gestione database (attività delle banche dati)

UNITÀ DI COMPETENZA - Raccolta dati ed applicazione di metodi e tecniche di data mining
RISULTATO ATTESO DALL'ESERCIZIO DELLA COMPETENZA Eeguire, sulla base delle esigenze di analisi e delle sorgenti informative disponibili, operazioni di raccolta dati, loro organizzazione e data mining, individuando ed applicando metodi e tecniche matematiche e statistiche
LIVELLO E.q.f.: 6

CONOSCENZE

- Caratteristiche della *business intelligence*
- Principi della *business intelligence* applicati all'economia aziendale
- Basi di dati relazionali (RDBMS) e No-SQL
- Linguaggio SQL e strumenti di interazione con basi dati No-SQL
- Metodi e tecniche di Data Gathering, Transformation e Configuration
- Strumenti di memorizzazione e recupero dati
- Metodi e tecniche di data mining (p.e. *cluster analysis, regression analysis, classification analysis, anomaly detection analysis, intrusion detection, association rule learning, decision tree, neural networks, rule induction, data warehouse*)
- Metodi di analisi di informazioni non strutturate
- Normativa vigente in materia di privacy
- Diritto in materia di proprietà intellettuale dei dati
- Lingua inglese di settore

ABILITA'

- Definire, sulla base delle esigenze aziendali e dei sistemi informativi, i fabbisogni di data mining
- Identificare e caratterizzare le sorgenti informative
- Svolgere operazioni di Data Gathering, Data Transformation e Data Configuration, sulla base di caratteristiche ed eterogeneità delle fonti grezze, utilizzando il linguaggio SQL e altri strumenti di interazione con basi di dati
- Selezionare i dati, in ragione degli obiettivi di analisi
- Preparare i dati selezionati, eliminando possibili errori e stabilendo dei meccanismi di comportamento in caso di dati mancanti, anche attraverso approcci campionari
- Applicare tecniche di data mining, anche adattando gli algoritmi sulla base delle esigenze di analisi
- Tracciare le operazioni compiute e valutarne gli esiti, anche in vista del consolidamento del processo
- Garantire proprietà intellettuale e rispetto della privacy

INDICATORI DI VALUTAZIONE DEL POSSESSO DELLA COMPETENZA

Sulla base di indicazioni relative ad un fabbisogno aziendale e di un set di sorgenti informative, individuare ed applicare tecniche di data mining, impostando preliminarmente le operazioni di raccolta e preparazione dati

PRESTAZIONE MINIMA ATTESA IN ESITO ALLA VALUTAZIONE

Per almeno una tipologia di fabbisogno, motivata individuazione ed applicazione dell'insieme delle tecniche coerenti di preparazione dati e di almeno due tecniche di data mining

MODALITÀ DI VALUTAZIONE DEL POSSESSO DELLA COMPETENZA

Audizione, colloquio tecnico e/o prova prestazionale

UNITÀ DI COMPETENZA - Realizzazione di analisi complesse e su big data**RISULTATO ATTESO DALL'ESERCIZIO DELLA COMPETENZA**

Realizzazione di analisi avanzate di data set (anche di notevoli dimensioni e di big data), creando modelli, algoritmi e metodi di raccolta, standardizzazione ed analisi dati, realizzando reportistica a supporto dei processi decisionali

LIVELLO E.q.f.: 6**CONOSCENZE**

- Linguaggio SQL e strumenti di interazione con basi dati No-SQL
- Linguaggi di programmazione (R, Python, ...) orientati ai data analytics
- Problematiche e specificità dei big data: dimensione e destrutturazione dei data set
- Metodi e tecniche di data curation, data quality control, data integrity and interoperability
- Metodi e tecniche di data mining
- Modelli analitici matematici e statistici, descrittivi e inferenziali: Forecasting modeling, Natural Language Processing, Clustering modeling, Classification modeling
- Elementi di machine learning e deep learning
- Tipologie di analisi: Descriptive Analytics; Diagnostic Analytics; Predictive Analytics; Prescriptive Analytics
- Metodi, tecniche e strumenti di Data visualization
- Normativa vigente in materia di privacy
- Lingua inglese di settore

ABILITA'

- Standardizzare ed ottimizzare i processi di gestione ed aggiornamento dei dati, implementando procedure di data curation, data quality control, data integrity and interoperability
- Creare modelli, algoritmi, strumenti per la raccolta e l'analisi di big data, sulla base delle esigenze aziendali, utilizzando linguaggi di programmazione
- Sviluppare, sulla base delle esigenze, algoritmi di machine learning
- Utilizzare big data analytics platforms
- Rappresentare gli esiti dell'analisi, utilizzando tecniche e strumenti di data visualization

INDICATORI DI VALUTAZIONE DEL POSSESSO DELLA COMPETENZA

Sulla base di indicazioni relative ad un fabbisogno aziendale e di un set di sorgenti informative a natura di big data, individuare ed impostare la modellistica di analisi, l'approccio al machine learning ed al data visualization, individuando le problematiche relative all'implementazione di procedure di data curation, data quality control, data integrity and interoperability

PRESTAZIONE MINIMA ATTESA IN ESITO ALLA VALUTAZIONE

Per almeno una tipologia di fabbisogno, motivata individuazione di almeno due tipologie di analisi e conseguente scelta ed impostazione di metodi e strumenti, incluso il ricorso al machine learning e data visualization, descrivendo l'approccio alla definizione delle procedure di data curation, data quality control, data integrity and interoperability

MODALITÀ DI VALUTAZIONE DEL POSSESSO DELLA COMPETENZA

Audizione, colloquio tecnico e/o prova prestazionale

UNITÀ DI COMPETENZA - Sviluppo dei processi di data analytic**RISULTATO ATTESO DALL'ESERCIZIO DELLA COMPETENZA**

Pianificare e coordinare i processi di data analytic, assicurando il corretto feeding dei livelli strategici e decisionali

LIVELLO E.q.f.: 6

CONOSCENZE

- Principi della *business intelligence* applicati all'economia aziendale
- Modelli di organizzazione dei processi di data analytic (data-driven; centralizzato, matriciale, ...)
- Metodi di trasferimento degli esiti delle analisi, ai livelli aziendali in cui avviene l'elaborazione delle strategie di business
- Lingua inglese di settore

ABILITA'

- Pianificare i processi di data analytic di una organizzazione, sulla base delle proprie esigenze strategiche
- Definire ed implementare il modello organizzativo dei processi di data analytic
- Trasferire gli esiti delle analisi, ai livelli aziendali in cui avviene l'elaborazione delle strategie di business
- Monitorare i processi di data analytic ed il loro impatto sul business

INDICATORI DI VALUTAZIONE DEL POSSESSO DELLA COMPETENZA

Sulla base di indicazioni relative ad una azienda ed alle sue esigenze strategiche, impostare logica ed organizzazione dei processi di data analytic, individuando le modalità di trasferimento ai decisori dei relativi esiti

PRESTAZIONE MINIMA ATTESA IN ESITO ALLA VALUTAZIONE

Per almeno una tipologia di azienda, motivata rappresentazione dell'approccio all'organizzazione dei processi di data analytic e di gestione del trasferimento dei loro esiti applicativi, ai soggetti con ruolo decisionale

MODALITÀ DI VALUTAZIONE DEL POSSESSO DELLA COMPETENZA

Audizione, colloquio tecnico e/o prova prestazionale