

Progetto Da.Re.

Intellectual Output 2: Moduli di Data Science e Risorse Aperte per la Formazione

Sintesi e traduzione in lingua Italiana

La nostra ricerca ha identificato uno spazio non soddisfatto dalle università o da boot camp nell'educazione legata alla scienza dei dati. Tale spazio è costituito da quello che chiamiamo "Bridge Person" (persona ponte): colui che combina la conoscenza del business di una specifica organizzazione con una sufficiente competenza e comprensione della scienza dei dati, tale da collegare persone "non tecniche" addette al business con soggetti altamente formati per fornire valore aggiunto al business.

Crediamo che il ruolo di Bridge Person e di Responsabile dell'Analisi Dati identificati dal progetto Da.Re. possano fornire un significativo impatto nel miglioramento dell'applicazione della scienza dei dati nelle organizzazioni europee.

La nostra soluzione permette di abilitare un'ulteriore tipologia di formazione in Data Science per:

- Soggetti che hanno un'approfondita conoscenza dell'attività delle organizzazioni ma scarsa competenza legata alla scienza dei dati (questi potranno diventare responsabili dell'analisi dati dopo un relativo periodo di formazione);
- Studenti laureati che vogliono lavorare come Data Scientist nel business e che siano motivati ad acquisire maggiori competenze di valore per le loro future posizioni lavorative (questi potranno aspirare alla figura di responsabile dell'analisi dati a seguito di un'adeguata formazione);
- Soggetti di esperienza che hanno poco tempo ma vogliono conoscere il valore aggiunto che la scienza dei dati può fornire al loro business e come poter approcciare l'uso di questa scienza.

Uno scorcio sul programma di Da.Re. incluso nell'I.O.2 e nei suoi allegati

Il programma di Da.Re. prevede 2 parti: 80 ore di formazione online seguite da 70 ore di corso in presenza.

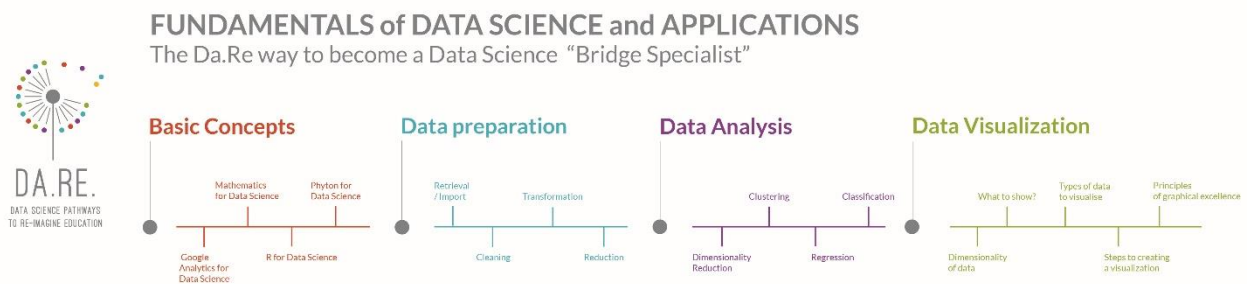
L'idea è di fornire agli studenti attraverso la formazione online la conoscenza tecnica utile e le competenze necessarie ad affrontare la parte in presenza del corso.

Combinando le due formule formative Da.Re. riesce a fornire i migliori elementi contenuti nei MOOC e nei BOOTCAMP al fine di fornire nuove, utili e sostenibili forme di educazione da diffondere in Europa.

I casi studio fornite per la parte in presenza del corso sono:

1. L'impronta energetica di una micro-rete elettrica;
2. L'analisi stazionaria per sistemi industriali
3. La gestione di progetti di comunicazione aziendali
4. L'analisi di serie temporali per finalità biomediche.

La parte online del corso è rappresentata dalla seguente immagine:



Il corso è strutturato in 5 differenti moduli, 4 teorici ed uno pratico. Quest'ultimo è costituito dai casi studio inclusi nel I.O.2 come allegati.

La parte teorica contenuta nel testo del I.O.2 è la seguente:

BASIC CONCEPTS (Concetti di Base)

In questo modulo gli studenti impareranno i concetti fondamentali legati alla scienza dei dati. Il corso inizia con l'apprendimento legato agli Analytics di Google ed a come questi possano fornire una grande varietà di informazioni per l'uso di siti internet, sugli utilizzatori ed i loro comportamenti in rete. A seguire è possibile avere un'introduzione all'uso di Python e delle librerie standard legate all'analisi dati, tra cui NumPy, Pandas, Scikit-learn e Matplotlib. Inoltre, saranno fornite le competenze utili all'uso dei pacchetti di R che forniscono una grande quantità di metodi per l'analisi statistica dei dati. Il modulo si concluderà con una panoramica sulle competenze matematiche utili per la data science.

DATA PREPARATION (Preparazione dei Dati)

La preparazione dei dati ha un ruolo critico nella vita del Data Scientist. All'interno di questo modulo verranno forniti gli aspetti chiave per la preparazione dei dati al fine della successiva analisi, alcuni saranno affiancati ad esempi reali. In tal modo, saranno trattati argomenti quali il data loading, il sampling, la feature extraction e la trasformata di Fourier. Alla fine del modulo ci si aspetta che lo studente sia in grado di preparare con successo i dati per un'analisi efficace.

DATA ANALYSIS (Analisi dei Dati)

L'analisi dati rappresenta il modulo di accesso per il Machine Learning (ML). In questo modulo si affronterà il ML nelle sue diverse sfaccettature, quali la riduzione della dimensione, la classificazione, la predizione e la clusterizzazione. Queste tecniche sono presentate in maniera teorica mantenendo tuttavia un collegamento diretto con esempi pratici; vengono inoltre incluse implementazioni attraverso l'uso di Python. Alla fine del modulo lo studente saprà applicare i vari paradigmi per l'analisi dati ed interpretare i risultati che ne seguono.

DATA VISUALIZATION (Visualizzazione dei Dati)

Cosa fare una volta ottenuti i dati? In questo modulo gli studenti si focalizzeranno sulle possibili vie per la visualizzazione dei dati, utili a rispondere a questa domanda. In particolare, saranno apprese le modalità di visualizzazione dei dati utili per l'acquisizione di dettagli da fornire agli interlocutori. Saranno introdotte le teorie utili ad effettuare presentazioni efficaci. Saranno inoltre effettuati esercizi per l'utilizzo di R in casi pratici utili a differenti tipologie di visualizzazione.

Background e conoscenze pregresse necessarie al corso

- Livello di educazione 6 o superiore (Laurea Triennale)
- Livello di conoscenze matematiche per la risoluzione di equazioni e per la rappresentazione di curve e grafici
- Capacità di scrivere report e illustrare immagini
- Capacità di effettuare ricerche e sintetizzare informazioni
- Interesse sull'impatto dei dati nelle attività organizzative
- Capacità di organizzare il tempo e lo studio
- Capacità di lavorare in team.

Risultati di apprendimento

I risultati attesi dallo studio dell'I.O.2 sono i seguenti:

- La conoscenza del ciclo di vita del dato
 1. Preparazione
 2. Analisi
 3. Visualizzazione
- Pratica ed abilità nella risoluzione di problemi realmente sperimentabili nelle organizzazioni
- La capacità di andare oltre il ciclo di vita del dato al fine di fornire valore aggiunto per le organizzazioni
- La capacità di organizzare e revisionare il ciclo di vita del dato nell'organizzazione di appartenenza, identificando e selezionando le competenze esistenti e non esistenti, creando team e strutture di lavoro per affrontare il seguente processo
 1. Identificazione del problema
 2. Preparazione dei dati necessari
 3. Analisi dei dati
 4. Visualizzazione dei dati
 5. Risoluzione della problematica.



Tutti i contenuti dei 5 moduli del I.O.2 sono disponibili nella pagina di download del sito di Da.Re.