

In collaborazione con:

wyscout



Programma

Il corso è articolato su **5 giornate consecutive di 8 ore ciascuna per un totale di 40 ore** di lezioni frontali, in cui si alterneranno insegnamenti teorici ed esercitazioni pratiche, *guest lecture*, project work e laboratori; i partecipanti saranno guidati da docenti di alto livello all'apprendimento di tecniche utili all'organizzazione e comprensione dei big data e in particolare all'elaborazione di modelli per la match analysis e lo scouting nel calcio.

Il programma didattico è articolato intorno alle tematiche esposte di seguito:

21 Settembre 2020 - **Formazione erogata da**



GIORNO 1: Data Visualization con Tableau

- Introduzione dati Wyscout & Setting Up Connections and Data Sources
- Area di lavoro
- Analisi temporali
- Correlazioni fra dati
- Mappare dati sul campo (Shoot analysis, Passing Path)
- Calcoli
- % di Contributo
- Creazione di cruscotti interattivi

22 Settembre 2020 - **Formazione erogata da**



GIORNO 2. Introduzione ai dati di Wyscout

- WEB API: protocollo HTTP, CRUD e formato JSON
- Cosa sono le API di Wyscout e come funzionano
- Utilizzo libreria requests per raccolta e parsing di risorse remote
- Intro a Pandas e Pandas DataFrame
- Importare/esportare dati (JSON e CSV)
- Manipolazione dei Pandas DataFrame
- Operazioni statistiche
- Gestione di valori nulli e outliers
- Esplorazione di variabili continue e categoriche
- Raggruppamenti: aggregazione di dati utilizzando DataFrame

Guest Lecture, **Wyscout**

23 Settembre 2020 - **Formazione erogata da** PlayeRank

GIORNO 3. Introduzione ai dati di Wyscout

Conoscere il dato su cui si lavora è il requisito principale per utilizzarlo al meglio. E' la prima fase nei processi di Data Science, chiamata Data Understanding. I partecipanti avranno quindi una panoramica sui dati forniti da wyscout e sul come filtrarli, aggregarli, analizzarne le distribuzioni. La lezione sarà interattiva: ai partecipanti saranno forniti i Jupyter notebook con cui seguire, passo-passo, le operazioni illustrate.

Argomenti affrontati:

- Parte 1
 - formato dei documenti *competition, player, team*
 - formato di un documento *match*
 - formato di un documento *event*
- Parte 2
 - aggregazioni di base su documenti event (numero passaggi, accuratezza, ecc.)
 - statistiche e visualizzazioni di base (posizione degli eventi, heatmap, ecc.)

Match data analysis

Nella match analysis esistono già metriche consolidate, fornite dai principali software ed utilizzate già dagli analisti nel loro lavoro quotidiano. Tuttavia,

l'analista e lo staff tecnico di cui fa parte hanno anche la loro particolare visione dei principi di sviluppo del gioco da andare a verificare. Ai partecipanti verrà illustrato come, partendo dai dati "grezzi", costruire le metriche principali e più conosciute.

Argomenti affrontati:

- Definizione, ricostruzione e analisi della rete di passaggi
 - Metriche di performance basate sulla rete di passaggi
 - Calcolo pericolosità e indice di invasione
 - Calcolo "expected goals"
-

24 Settembre 2020 - **Formazione erogata da** 

GIORNO 4. Player scouting

Lo scouting è uno degli utilizzi principali dei soccer logs, ovvero i dati relativi ad ogni evento sulla palla registrato in campo. La peculiarità principale di questo tipo di dati è infatti la sua estensione: non solo i principali campionati italiani, ma anche le seconde divisioni, i campionati primavera ed i campionati "minori" vengono coperti dalla raccolta dati, rendendo questo tipo di dato ottimo per la ricerca dei calciatori. La performance di un calciatore, nel nostro ambito, è un insieme di eventi osservati durante una partita disputata dal calciatore. Ai partecipanti verrà illustrato l'algoritmo Playerank, che permette di sfruttare tecniche di intelligenza artificiale per confrontare le performance dei calciatori, espresse sotto forma di eventi registrati sul campo. Playerank fornisce inoltre all'analista la possibilità di personalizzare l'algoritmo e di modellare e validare la propria ipotesi di valutazione della performance.

Argomenti affrontati

- valutazione delle performance con l'algoritmo PlayeRank
- ricerca nel database dei calciatori e confronto dell'evoluzione della performance
- personalizzazione della valutazione attraverso cambio dei parametri dell'algoritmo
- calcolo della similarità tra calciatori

Project Work

Suddivisione dei partecipanti al corso in gruppi di lavoro guidati dai docenti di Top-IX e PlayeRank

Guest Lecture, *Top Club Serie A*

25 Settembre 2020 - **Formazione erogata da**

PlayeRank

top-ix
WE DO CONNECTIONS

GIORNO 5**Project Work**

Suddivisione dei partecipanti al corso in gruppi di lavoro guidati dai docenti di Top-IX e PlayeRank

Esposizione e condivisione Project Work

Al termine della giornata ci sarà la consegna della certificazione per “*Football Data Analyst Wylab Top-IX in collaborazione con Wyscout*”.