



Serious Game Data Engineer

DS050

Durée: 1 jour

Public :

Chefs de projet, architectes, data-engineer

Objectifs :

Pour plonger de manière ludique mais toujours très dense, dans le domaine des data-sciences, et valider ou affiner des connaissances de data-engineer, et d'architecture bigdata.

Connaissances préalables nécessaires :

Connaissances BigData et IA (Bases de données NoSQL, Cassandra, elasticStack ,logstash, Machine Learning, Spark SQL, Dataiku DSS)

Programme :

Le scénario

Des candidats "Data Engineer" doivent être mis à l'épreuve par une société de data-sciences, afin de recruter le meilleur d'entre eux.

Les missions à réaliser

extraction de données depuis une base Cassandra, (extraits de Movies Dataset)
préparation des données pour l'entraînement d'un modèle machine learning de prédiction sur Dataiku,
analyse de ces données avec Spark-sql,
création d'un pipeline avec ElasticSearch/logstash,
prédiction sur le nouveau jeu de données généré.

La méthode

Simulation d'un cas d'étude, avec un travail collaboratif sur des données réelles, accessibles en opendata, et des labs techniques.
Épreuves personnelles et épreuves en commun permettent de valider les connaissances et d'échanger entre participants, tout en bénéficiant du soutien et des explications complémentaires du formateur sur les thèmes proposés

Les jeux

puzzles, Battle d'architecture, les points de faiblesse, etc...



— Phirio —

Le debrief

Retour des travaux, bilan des points individuels et classement des joueurs.
Retour d'expérience des participants