

BimBamJob - Test Technique

Data Scientist

Énoncé

Dans le cadre de votre test technique, vous devez proposer une solution à un challenge de data science sur la recommandation de poste. Il s'agit de recommander des emplois (*jobs.csv*) à des candidats (*users.csv*) en fonction des caractéristiques de chacun.

Pour cela, vous disposez de l'historique des candidatures (*applications.csv*) de candidats à un emploi ainsi que de données historiques des précédents postes occupés par les candidats (*users_history.csv*).

Vous devez recommander des emplois (*JobID*) pour les candidats issus du fichier *test_users.csv*.

Instructions

Vous êtes libre d'utiliser les données que vous souhaitez et de définir votre métrique d'évaluation pour ce problème. L'objectif de l'exercice **n'est pas d'évaluer la performance de la modélisation** mais d'évaluer la méthodologie et la réflexion mise en œuvre sur ce type de problème.

Cette étude ne devrait pas prendre plus de 4h, il n'est donc pas nécessaire d'investir trop sur la partie modélisation, néanmoins les idées pourront être discutées lors de l'entretien.

Lors de l'entretien technique, vous aurez l'occasion de présenter votre travail et le code associé.

Retour attendu

- Code python associé à la modélisation ou/et l'exploration des données
- Instructions d'installation ou d'utilisation pour rejouer le code

Données

jobs.csv :

- **JobID** : Id de l'emploi
- **Title** : Intitulé
- **Description** : Description du poste
- **Requirements** : Prérequis pour le poste
- **City, State, Country, Zip5** : Informations géographiques
- **StartDate, EndDate** : Date de début et fin de visibilité du poste sur le site d'annonce

users.csv :

- **UserID** : Id du candidat
- **City, State, Country, ZipCode** : Informations géographiques
- **DegreeType, Major, GraduationDate** : Information sur le diplôme du candidat
- **WorkHistoryCount** : Nb d'expériences
- **TotalYearsExperience** : Expérience totale
- **CurrentlyEmployed** : En poste
- **ManagedOthers, ManagedHowMany** : Si le candidat a occupé un poste de management et combien de managés.

users_history.csv :

- **UserID** : Id du candidat
- **Sequence** : Ordre de l'emploi occupé
- **JobTitle** : Intitulé de l'emploi occupé

applications.csv :

- **UserID** : Id du candidat
- **ApplicationDate** : Date à laquelle le candidat a postulé
- **JobID** : Id de l'emploi

test_users.csv : Candidats de test sur lesquels faire des recommandations.