

## 1er Data Challenge de l'ISDM

### Confiez vos données à des étudiants pour les explorer

L'institut de science des données<sup>1</sup> (ISDM) de Montpellier et la start-up Trustii<sup>2</sup> s'associent pour proposer le 1<sup>er</sup> data challenge de l'ISDM.

Vous souhaitez construire un modèle prédictif par machine learning sur vos données ? Vous souhaiteriez explorer les possibilités du machine learning pour vos données ? Le data challenge de l'ISDM vous offre l'opportunité de diffuser ces données et la problématique associée auprès des étudiants de l'Université de Montpellier. A l'issue de ce challenge, le modèle le plus performant sera mis à votre disposition, parmi les propositions de l'ensemble des étudiants participants.

Les données seront hébergées par la plateforme Trustii (<https://www.trustii.io>), à laquelle auront accès les étudiants candidats durant le challenge.

#### **Quelles données et quels problèmes peuvent être proposés dans le cadre de ce challenge ?**

Les méthodes visées sont de type "supervisées". Les données devront être présentées sous la forme de fichiers .csv intégrant des attributs de description et la classe d'appartenance.

Exemple de la célèbre base iris :

```
sepal_length,sepal_width,petal_length,petal_width,species  
5.1,3.5,1.4,0.2,1  
7,3.2,4.7,1.4,2  
6.4,3.2,4.5,1.5,2  
7.7,2.6,6.9,2.3,3
```

Ces données serviront à entraîner des algorithmes permettant de reconnaître et prédire la catégorie (dans notre exemple, la variété d'iris donnée par 1, 2 ou 3).

Les données doivent être préparées pour être interprétées par machine learning. Les champs du fichier CSV doivent être des nombres réels et non vides. Trustii pourra aider à la vérification du format et de la cohérence des données et orienter le cas échéant vers des outils de transformation de vos données.

---

<sup>1</sup> <https://isdumontpellier.fr/>

<sup>2</sup> <https://www.trustii.io/>

### Présentation de l'ISDM

L'institut de Science des Données de Montpellier (ISDM) est une structure de l'Université de Montpellier (UM) travaillant en articulation avec les partenaires de MUSE visant à :

- Structurer l'écosystème montpelliérain en science des données
- En faire connaître le potentiel toutes communautés confondues
- Favoriser et diffuser l'usage de la science des données et promouvant le mouvement de la science ouverte.

En ce sens, en lien avec tous les acteurs impliqués, l'ISDM est un relai de la dynamique du site, il contribue à l'accompagnement des projets, aux événements (hackathons, séminaires, ateliers,...) et à la création et diffusion de ressources.

L'ISDM est le porteur du DU "Scientific Data Management" (SDM).

**Informations et contact :** <https://isdm.umontpellier.fr/> -- [isdm-challenges@umontpellier.fr](mailto:isdm-challenges@umontpellier.fr)

### Présentation de Trustii

Trustii est une plate-forme collaborative dans le domaine de la science des données. Elle permet aux organisations propriétaires de données de créer des modèles de machine learning prêt à être déployés.

Les propriétaires de données soumettent leurs jeux de données à Trustii sur la plateforme en ligne. Trustii permet de proposer un modèle de machine learning de bonne qualité en agrégeant l'ensemble des modèles proposés par les data scientists inscrits sur la plateforme.

Les propriétaires des jeux de données peuvent ensuite télécharger ou exécuter le modèle directement sur la plate-forme.

**Informations et contact :** <https://www.trustii.io/> -- [challenges@trustii.io](mailto:challenges@trustii.io)

### Calendrier de l'appel

Lancement de l'appel	20/11/2020
Date limite de proposition de jeu de données	04/01/2021
Publication des jeux de données Et Lancement des accès étudiants	25/01/2021
Date limite de proposition de modèle	15/02/2021
Déclaration des résultats et remise des prix	19/02/2021

## 1<sup>er</sup> data challenge de l'ISDM - proposition de jeu de données

Formulaire à remplir en ligne avant le 04/01/2020 :

<https://questionnaire.umontpellier.fr/index.php/392195?lang=fr>

pour toute question : [isdm-challenges@umontpellier.fr](mailto:isdm-challenges@umontpellier.fr)

Nom et Prénom :

e-mail :

Organisation de rattachement (université, organisme, entreprise, collectivité, ...) :

### Descriptif du jeu de données

### Descriptif de la problématique et du modèle attendu

Exemples : Je souhaiterais prédire le coût d'une maison à Montpellier en fonction de sa localisation, surface...

### Engagement

En confiant mon jeu de données pour le data challenge de l'ISDM-Trustii, je certifie avoir pris connaissance des conditions et engagements ci-dessous.

- Je m'engage à accepter les conditions générales de la plateforme Trustii.
- Je certifie que les données sont conformes à tous les règlements en vigueur, et notamment le RGPD.
- Je certifie que les données ne font l'objet d'aucune clause de propriété interdisant leur partage.
- Je reconnais avoir pris connaissance du fait que les données seront hébergées sur la plateforme de Trustii.
- Je reconnais avoir pris connaissance que mes données seront traitées par un panel d'étudiants inscrits à l'Université de Montpellier accédant à la plateforme Trustii.
- Je reconnais avoir pris connaissance du fait que je ne pourrai pas interdire l'accès à mes données à un quelconque étudiant dûment inscrit à l'Université de Montpellier.
- Je reconnais avoir pris connaissance du fait que les étudiants participants gardent la propriété intellectuelle des modèles dont ils seront l'auteur mais que lesdits modèles seront accessibles au sein de la plateforme Trustii.
- Je reconnais avoir pris connaissance du fait que les propriétaires des données peuvent utiliser les modèles pour un usage non-commercial et que toute utilisation des modèles dans un cadre commercial doit faire l'objet d'un contrat commercial avec Trustii prévoyant une rémunération des étudiants data scientists ainsi qu'une utilisation libre des modèles.

### Proposition de lots pour les étudiants

Si vous le souhaitez, vous pouvez proposer des lots récompensant les étudiants ayant fourni des modèles performants. Le cas échéant, les lots que vous proposez seront les suivants :

- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_

**Nom, date et signature**