

# Intelligence Artificielle 3

## Cours 1 - Règlements de compétition

Steve Lévesque, Tous droits réservés © où applicables

# Table des matières

- 1 Introduction au site Kaggle
- 2 Notebooks Kaggle
- 3 Compétition - Comment y participer ?
- 4 Règlements de compétition
- 5 Compétition - Bonne chance !!!

# Introduction au site Kaggle

Kaggle est un forum/site pour la communauté des scientifiques de données très populaire pour ses compétitions, les ressources gratuites faites par une communauté d'experts et son aspect social collaboratif intéressant.

# Introduction au site Kaggle

On peut à partir du site Kaggle :

- Voir et participer à des compétitions
- Voir des Notebooks et datasets d'experts
- Créer des Notebooks, datasets, modèles, compétitions et plus encore
- Discuter avec enthousiastes des sciences de données avec la partie sociale du site (commentaires)
- Apprendre la science des données, l'AI, le Machine Learning, etc.
- Encourager et avoir une mentalité “open-source”
- Et plus encore !

# Notebooks Kaggle

**kaggle** Competitions Datasets Models Code Discussions Courses ...


Q Search

Sign in Register

## Level up with the largest AI & ML community

Join over 18M+ machine learners to share, stress test, and stay up-to-date on all the latest ML techniques and technologies. Discover a huge repository of community-published models, data & code for your next project.


Register with Google Register with Email



### Who's on Kaggle?


**Learners**

Dive into Kaggle courses, competitions & forums.




**Developers**

Leverage Kaggle's models, notebooks & datasets.



**Researchers**

Advance ML with our pre-trained model hub & competitions.



See more

Figure: <https://www.kaggle.com/>

# Notebooks Kaggle

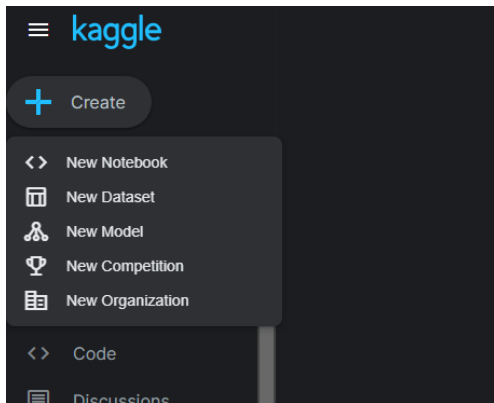


Figure: <https://www.kaggle.com/>

# Notebooks Kaggle

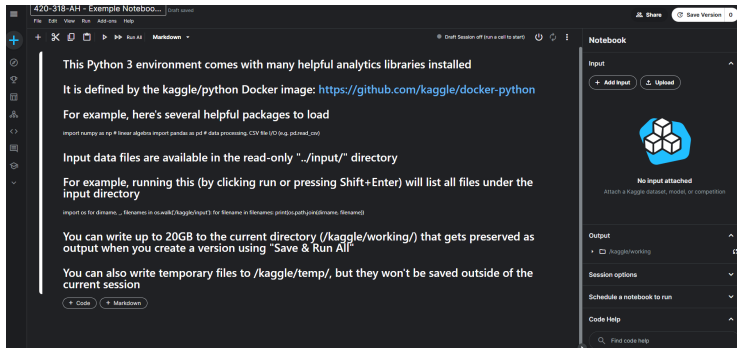


Figure: <https://www.kaggle.com/>

# Notebooks Kaggle

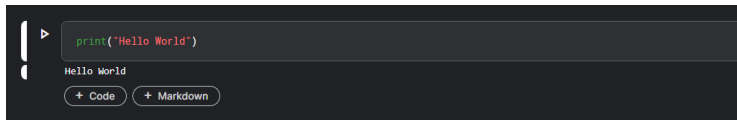


Figure: <https://www.kaggle.com/>



# Compétition - Comment y participer ?

Attention : Pour la compétition privée du cours, le professeur va vous inviter à l'aide de votre courriel Ahuntsic.

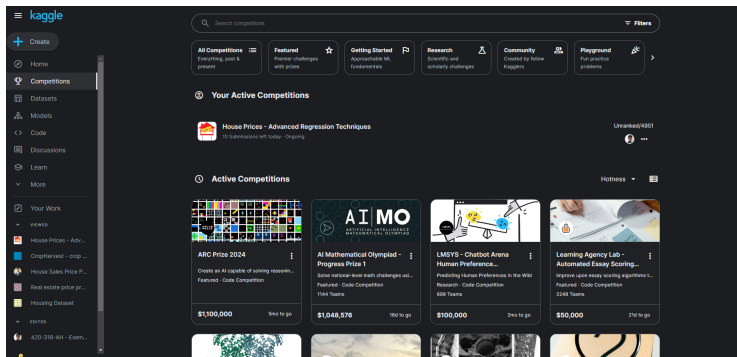


Figure: <https://www.kaggle.com/>

# Compétition - Comment y participer ?

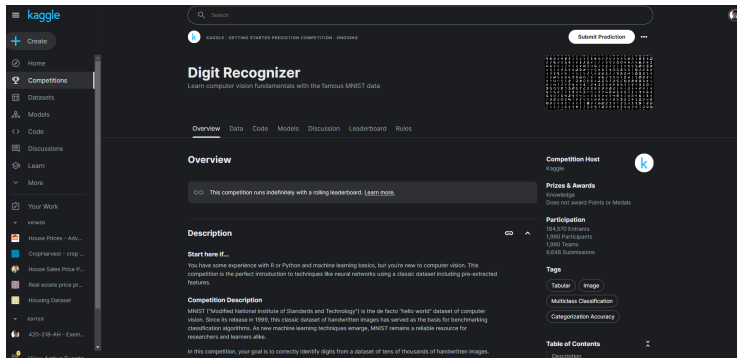
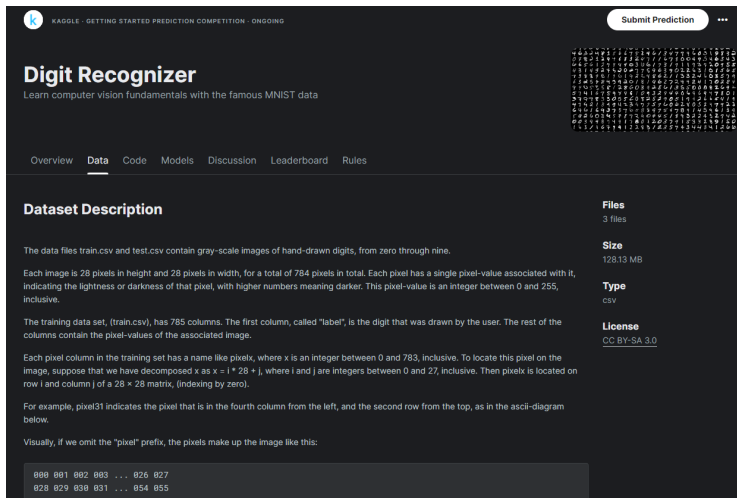


Figure: <https://www.kaggle.com/>

# Compétition - Comment y participer ?



**KAGGLE - GETTING STARTED PREDICTION COMPETITION - ONGOING**

## Digit Recognizer

Learn computer vision fundamentals with the famous MNIST data

Overview **Data** Code Models Discussion Leaderboard Rules

### Dataset Description

The data files train.csv and test.csv contain gray-scale images of hand-drawn digits, from zero through nine.

Each image is 28 pixels in height and 28 pixels in width, for a total of 784 pixels in total. Each pixel has a single pixel-value associated with it, indicating the lightness or darkness of that pixel, with higher numbers meaning darker. This pixel-value is an integer between 0 and 255, inclusive.

The training data set, (train.csv), has 785 columns. The first column, called "label", is the digit that was drawn by the user. The rest of the columns contain the pixel-values of the associated image.

Each pixel column in the training set has a name like pixel*x*, where *x* is an integer between 0 and 783, inclusive. To locate this pixel on the image, suppose that we have decomposed *x* as  $x = i * 28 + j$ , where *i* and *j* are integers between 0 and 27, inclusive. Then pixel*x* is located on row *i* and column *j* of a  $28 \times 28$  matrix, (indexing by zero).

For example, pixel31 indicates the pixel that is in the fourth column from the left, and the second row from the top, as in the ascii-diagram below.

Visually, if we omit the "pixel" prefix, the pixels make up the image like this:

```
000 001 002 003 ... 026 027
028 029 030 031 ... 054 055
056 057 058 059 ... 082 083
```

**Files**  
3 files

**Size**  
128.13 MB

**Type**  
CSV

**License**  
[CC BY-SA 3.0](#)

Figure: <https://www.kaggle.com/>

# Compétition - Comment y participer ?

The screenshot shows the Kaggle 'Digit Recognizer' competition page. At the top, there's a header with the Kaggle logo and the text 'KAGGLE · GETTING STARTED PREDICTION COMPETITION · ONGOING'. A 'Submit Prediction' button is in the top right. The main title 'Digit Recognizer' is prominently displayed, followed by the subtitle 'Learn computer vision fundamentals with the famous MNIST data'. Below this, there are tabs for 'Overview', 'Data', 'Code' (which is selected), 'Models', 'Discussion', 'Leaderboard', and 'Rules'. A 'Notebooks' section is visible, featuring a search bar and a list of notebooks. The notebooks are ranked by 'Most Votes'. The top notebook is 'Introduction to CNN Keras - 0.997 (top 6%)' by a user with a profile picture, updated 7mo ago, with 1060 comments and 6825 votes (Gold medal). Other notebooks include 'Pytorch Tutorial for Deep Learning Lovers' (1607 votes), 'Deep Neural Network Keras way' (1125 votes), and 'Knowledge Graph & NLP Tutorial-(BERT,spaCy,NLTK)' (1047 votes).

KAGGLE · GETTING STARTED PREDICTION COMPETITION · ONGOING

Submit Prediction

## Digit Recognizer

Learn computer vision fundamentals with the famous MNIST data

Overview Data **Code** Models Discussion Leaderboard Rules

### Notebooks

+ New Notebook

Search notebooks Filters

All Your Work Shared With You Bookmarks Most Votes

- Introduction to CNN Keras - 0.997 (top 6%)**  
Updated 7mo ago  
1060 comments · Digit Recognizer  
6825 votes Gold
- Pytorch Tutorial for Deep Learning Lovers**  
Updated 4y ago  
148 comments · Digit Recognizer  
1607 votes Gold
- Deep Neural Network Keras way**  
Updated 6y ago  
Score: 0.99028 · 206 comments · Digit Recognizer  
1125 votes Gold
- Knowledge Graph & NLP Tutorial-(BERT,spaCy,NLTK)**  
Updated 4y ago  
88 comments · Digit Recognizer  
1047 votes Gold

Figure: <https://www.kaggle.com/>

# Compétition - Comment y participer ?

The screenshot shows the 'Digit Recognizer' competition page on Kaggle, specifically the 'Discussion' tab. The page has a dark theme. At the top, there's a search bar and a 'Submit Prediction' button. Below the navigation tabs (Overview, Data, Code, Models, Discussion, Leaderboard, Rules), the 'Discussion' section is active. It includes a search bar for discussions, a 'Follow' button, and a 'New Topic' button. The discussion list is filtered by 'All' topics and sorted by 'Most Votes'. The list includes pinned topics and other discussions. Each topic entry shows a profile picture, the title, the author, the last comment time, and the number of comments.

Topic	Author	Last Comment	Comments
<b>Rolling Leaderboards</b>	Will Cukierski	4y ago by Darya Vozhdseva	137 comments
<b>Getting Started Competitions - 100% Public Leaderboard</b>	Will Cukierski	4y ago by Junming lao	25 comments
<b>This Competition has an Official Discord Channel</b>	Myles O'Neill	12d ago by Yousef Mohammad	22 comments
<b>How to score 97%, 98%, 99%, and 100%</b>	Chris Deotte	1d ago by Fahmida Chowdhury	260 comments
<b>Kaggle Scripts</b>	Ben Hamner	4y ago by Juan José García Urrea	28 comments
<b>You might be cheating without knowing it</b>	Samuel Corbin	4y ago by Shivaditta Vaidyanathan	14 comments

Figure: <https://www.kaggle.com/>

# Compétition - Comment y participer ?

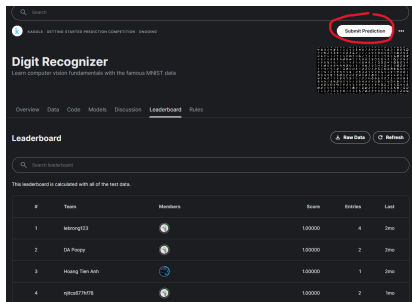


Figure: Remettre vos prédictions sur Kaggle.

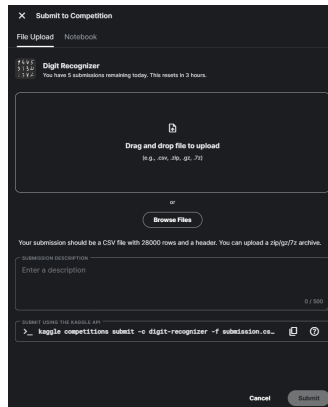


Figure: Un fichier .csv est standard pour les remises.

## **Tricher est du plagiat.** Que constitue tricher ?

Infractions spéciales de la compétition qui comptent comme du plagiat :

- Changer les “labels” manuellement à la main pour avoir un meilleur score dans la compétition lors de la soumission sur Kaggle
- Prendre et utiliser le résultats d'une autre équipe lors de la soumission sur Kaggle

# Règlements de compétition

**Tricher est du plagiat.** Que constitue tricher ?

Rappels non exhaustifs des infractions académiques traditionnelles qui s'applique pour cette compétition :

- Utiliser le code ou des méthodologies de l'Internet ou d'ouvrages sans faire de citations appropriées
- Utiliser une partie ou en entier de peu importe (rapport, code, etc.) d'une autre équipe



# Règlements de compétition

**Tricher est du plagiat.** Que constitue tricher ?

Normalement, mélanger le “train set” et “test set” est considéré de la triche (“cheating”) dans la littérature puisque vous évaluez à partir de données déjà vue par le modèle lors de l’entraînement.

Ceci-dit, **SVP faites attention pour ne pas le faire dans le cadre de votre travail.** Par contre, même par malheur que vous oubliez, ce n’est pas compté comme du plagiat.

Ce sera compté comme une erreur importante dans le barème.

# Compétition - Bonne chance !!!

Une compétition privée pour le cours sera bientôt disponible.

**SVP créer le compte Kaggle à partir du courriel du Collège Ahuntsic pour que je puisse vous inviter.**

Veuillez vous référer à l'énoncé et au plan de cours pour plus de détails sur les pondérations, la date de remise de toutes les composantes (essais de compétition, code, rapport, présentation orale, etc.) ainsi que le nombre de personnes étudiantes nécessaires par équipe.

Bonne chance !!!

# Bibliographie

■ <https://www.kaggle.com/>