

# Intelligence Artificielle 2

## Cours 3 - Notebooks

Steve Lévesque, Tous droits réservés © où applicables

# Table des matières

- 1 Démarrer avec Google Colab
- 2 Chargement des données massives
- 3 Représentation des arbres et graphes dans des cellules graphiques (plots)
- 4 Création de graphiques (plots) pour la visualisation des données massives, analyses statistiques, et des métriques
- 5 Notebooks et de l'IDE dans un projet d'IA (professionnel ou académique)
  - Les Notebooks
  - L'IDE
  - Professionnel et académique, bon mélange ?

# Démarrer avec Google Colab - Introduction

Google Colab est un environnement cloud de Notebooks gratuit, facile d'utilisation et performant.

Il permet l'utilisation des modules, de manière générale, sans installations puisqu'ils le sont déjà dans le cloud.

Il faut un compte Google pour l'utiliser et sauvegarder les projets.

# Démarrer avec Google Colab - Avoir un compte Google

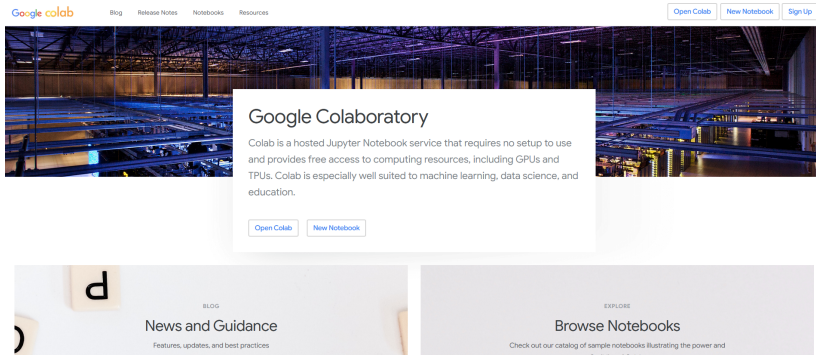


Figure: <https://colab.google/>

# Démarrer avec Google Colab - Notebook de bienvenu

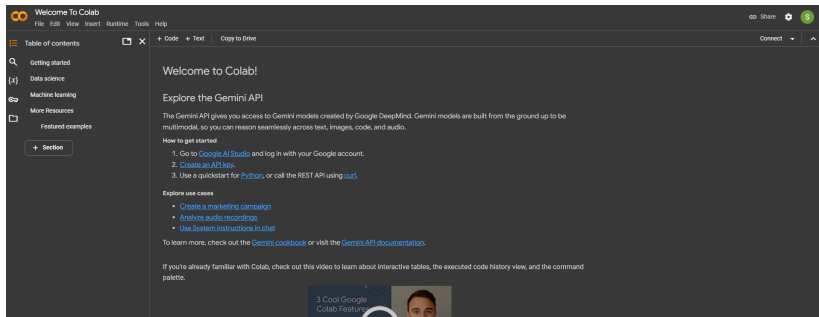


Figure: <https://colab.research.google.com/>

# Démarrer avec Google Colab - Notebook de bienvenu

Permet de coder ainsi que de voir le résultat isolé dans les cellules.

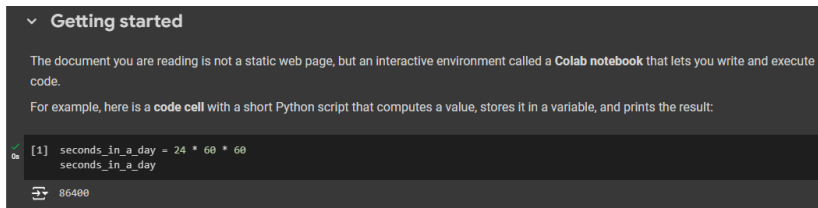


Figure: <https://colab.research.google.com/>

# Démarrer avec Google Colab - Notebook de bienvenu

La manière la plus simple d'exécuter du code d'une cellule est de peser sur le bouton "play".

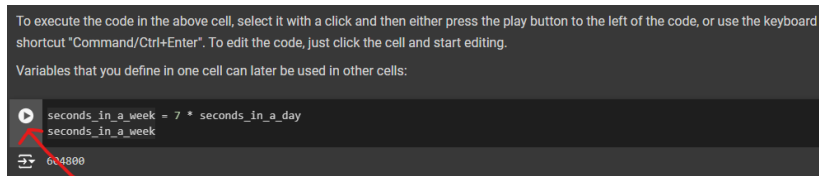


Figure: <https://colab.research.google.com/>

# Démarrer avec Google Colab - Notebook de bienvenu

Voici un graphique résultant du code d'une cellule. Généralement plus convivial à produire dans le Notebook que dans un IDE.

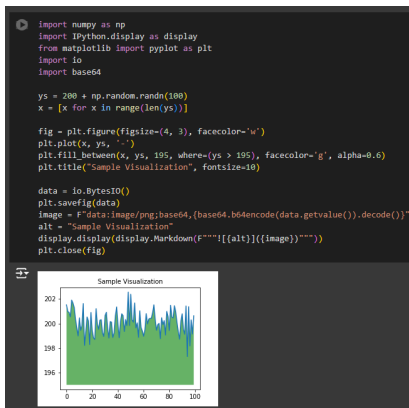


Figure: <https://colab.research.google.com/>



# Démarrer avec Google Colab - Notebook de bienvenu

“Disseminating” veut dire diffuser. La recherche est basé sur le partage ouvert.

NB : Le lien en caption sur la figure permet d'exécuter un exemple simple de tensorflow 2 sur un Notebook.

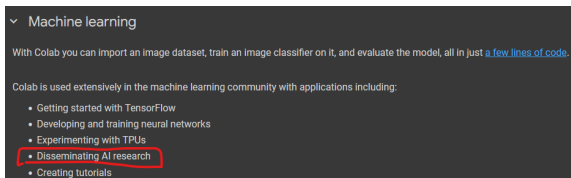


Figure: <https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/tutorials/quickstart/beginner.ipynb>

# Démarrer avec Google Colab - Notebook de bienvenu

Peu importe si c'est Google Colab (cloud) ou Jupyter Notebook (local sur l'ordinateur), les Notebooks sont souvent utilisés pour simplifier l'exécution surtout en lien avec le partage du savoir et les représentations graphiques.

Sur ce, il ne serait pas recommandé de faire de la programmation avancés d'applications robustes faite pour la production ou détenant du secret professionnel dans de simples Notebooks.

# Chargement des données massives

Des projets de Machine Learning sérieux ont généralement des centaines de milliers ou des millions d'entrées de données.

Il faut généralement beaucoup de données pour permettre au modèle d'apprendre adéquatement.

Il est possible d'utiliser Google Drive pour stocker les données massives et les accéder à partir des Notebooks de Google Colab.

# Chargement des données massives

```
1 from google.colab import drive
2
3 PREFIX = '/content/gdrive/'
4 drive.mount(PREFIX)
5
6 PATH = "MyDrive/"
7
8 # Read the file line by line until the end.
9 with open(PREFIX + PATH + 'test.txt') as file:
10     lines = file.readlines()
11     for line in lines:
12         print(line)
```

# Chargement des données massives

NB : Le fichier text contient seulement une ligne, mais l'import de fichiers massifs sont supportés de cette manière.

```
[1] from google.colab import drive

PREFIX = '/content/gdrive/'
drive.mount(PREFIX)

Drive already mounted at /content/gdrive/; to att

[2] PATH = "MyDrive/"

# Read the file line by line until the end.
with open(PREFIX + PATH + 'test.txt') as file:
    lines = file.readlines()
    for line in lines:
        print(line)
```

420-316-AH - Intelligence Artificielle 2

Figure: Utilisation du module google.colab pour monter le drive et accéder au fichiers comme un ordinateur traditionnel



Figure: Le fichier .txt dans le Google Drive avec le contenu respectif de la figure d'essai (à gauche)

# Représentation des arbres et graphes dans des cellules graphiques (plots)

Voir le Notebook disponible sur la plateforme d'apprentissage.

# Création de graphiques (plots) pour la visualisation des données massives, analyses statistiques, et des métriques

Voir le Notebook disponible sur la plateforme d'apprentissage.

# Notebooks et de l'IDE dans un projet d'IA (professionnel ou académique)

Les Notebooks et le code structuré par projet dans un IDE ont tous les deux une place importante dans un environnement de développement en Machine Learning.

Chacun détient des avantages et des défauts, et ils se complètent parfaitement.



# Projet d'IA - Les Notebooks

**Les Notebooks** permettent de représenter et partager de l'information (métriques, graphiques, etc.) très facilement et avec peu d'efforts de programmation et compilation. Aussi, la sauvegarde est sur le cloud et est accessible sur n'importe quel ordinateur capable de supporter un navigateur Web.

**Les projets et l'IDE** permettent de créer des applications robustes, complexes et avec des propriétés avancées (plus proche du OS, état clair des ressources utilisées, etc.). Il est plus difficile de faire des erreurs de programmation puisque la compilation est sujette à l'intégrité du projet au complet (non seulement une cellule) incluant les tests unitaires et le CI/CD.

# Projet d'IA - Professionnel et académique, bon mélange ?

Dans un cas général, pour le milieu professionnel ou recherche :

**Les Notebooks** démontrent les résultats d'exploration et les graphiques divers (composition des données, statistiques, métriques de fin d'exécution, etc.).

**Les projets et l'IDE** permettent d'intégrer le projet de Machine Learning en production (professionnel) ou d'avoir une sauvegarde documentée et robuste pour des extensions futures (recherche).

Possiblement, **les deux cas sont jumelés** puisqu'un projet de recherche peut être soutenu par l'industrie pour un besoin mutuellement bénéfique, et vice-versa, l'industrie pourrait vouloir faire de la recherche pour développer un tout nouveau produit novateur.

# Bibliographie

- <https://coin-or.github.io/pulp/main/index.html>
- [https://dav74.github.io/site\\_nsi\\_term/c9c/](https://dav74.github.io/site_nsi_term/c9c/)
- <https://www.geeksforgeeks.org/graph-theory-tutorial/>
- <https://www.geeksforgeeks.org/scipy-csgraph-compressed-sparse-graph/>