



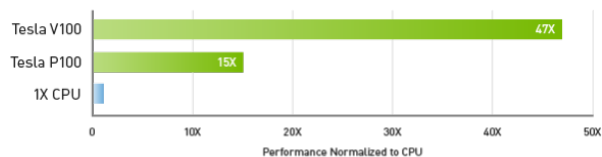
Introduction to Google Colab

2022.04.26
Jonghwa lee

Deep learning requires a lot of hardware

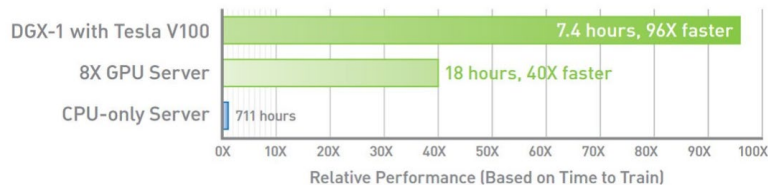
- To efficiently train a deep learning model, your hardware should be able to deal with **huge amount of data at high speed**
- CPUs are fast but their bandwidth is lower than GPUs
 - GPUs can perform a large number of matrix multiplications at a time
 - It allows you to build and train complex deep learning models on large dataset efficiently
- However, high-performance GPUs are expensive

47X Higher Throughput Than CPU Server on Deep Learning Inference



Workload: ResNet-50 | CPU: 1X Xeon E5-2690v4 @ 2.6 GHz | GPU: Add 1X Tesla P100 or V100

NVIDIA DGX-1 Delivers 96X Faster Training



Workload: ResNet50, 90 epochs to solution | CPU Server: Dual Xeon E5-2699 v4, 2.6GHz

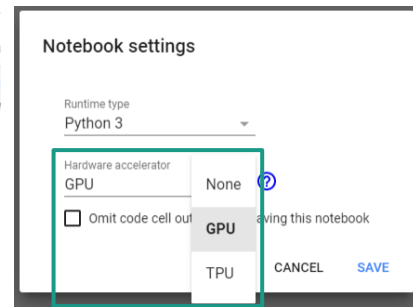
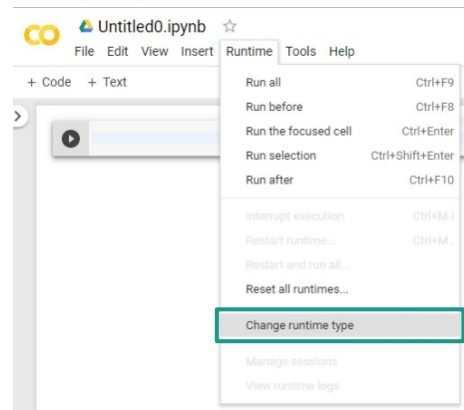
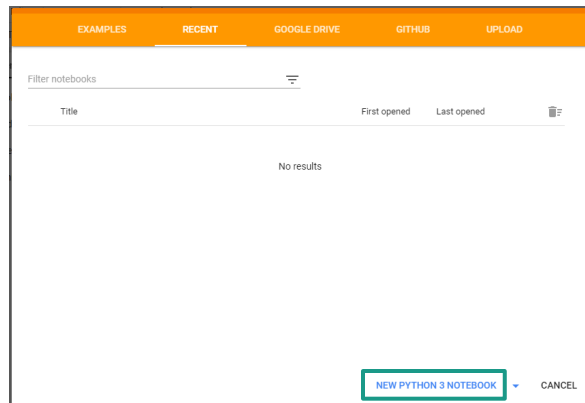


Google Colab to the rescue!

- Google Colab (<https://colab.research.google.com>) is a free Jupyter notebook environment that requires no setup and runs entirely in the cloud
- It currently offers the computing services of an NVIDIA Tesla K80 GPU **for free**
- You can use the computing services for a maximum of 12 hours at a time
 - Basically, you are connected to a virtual machine that lives for a maximum of 12 hours at a time
 - All data including model parameters that aren't saved to the Google Drive before this period will be lost

Getting Started

- Go to <https://colab.research.google.com>
- Create a new Python 3 notebook
- From the 'Runtime > Change runtime type' menu, assign a hardware accelerator to your notebook



Notebook

- Notebook is a **list of cells** which contain either **explanatory text** or **executable code and its output**
- By clicking a cell, you can select the cell that you want to working with
- Text cells help you explain your notebook and code cells you wrote

Text cells use markdown syntax. You can also add math to text cells using LaTeX. Just place the statement within a pair of \$ signs.
For example `$$\sqrt{3x-1}+(1+x)^2$` becomes $\sqrt{3x-1} + (1+x)^2$.

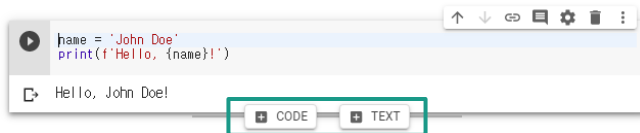
- Code cells contain executable Python code and its output

```
name = 'John Doe'
print(f'Hello, {name}!')
```

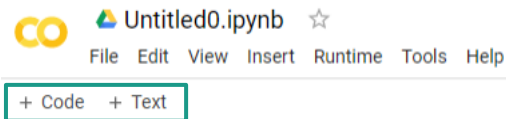
Hello, John Doe!

Working with Cells

- You can add new code or text cells using the **CODE** or **+TEXT** buttons

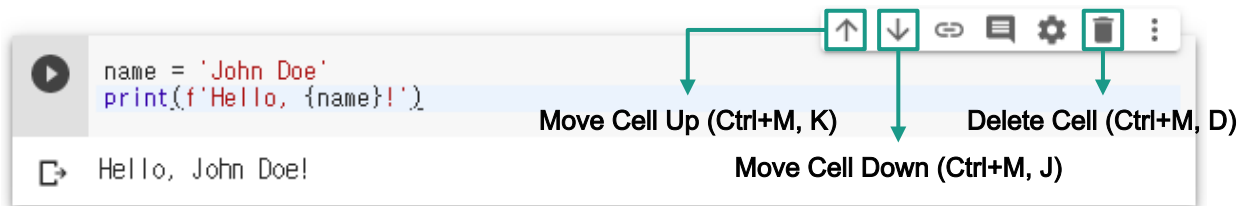


Buttons that appear when you hover between cells



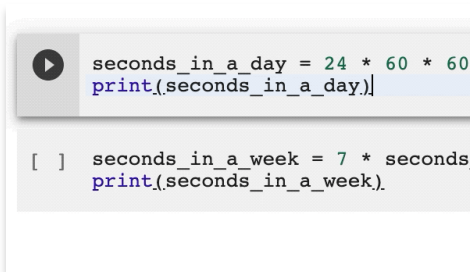
Buttons in the notebook toolbar

- By selecting a cell, you can move up, move down or delete the cell



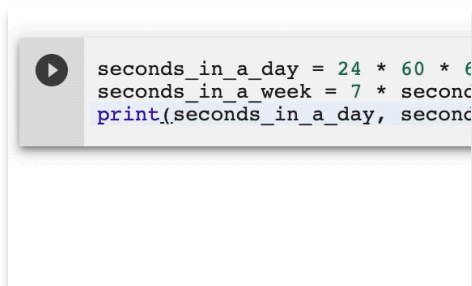
Working with Cells

- You can execute a selected code cell in the following ways:
 - Click the **Play button** in the left of the cell
 - Type **Ctrl+Enter** to run the cell in place
 - Type **Shift+Enter** to run the cell and move focus to the next cell
 - Type **Alt+Enter** to run the cell and insert a new code cell immediately below it



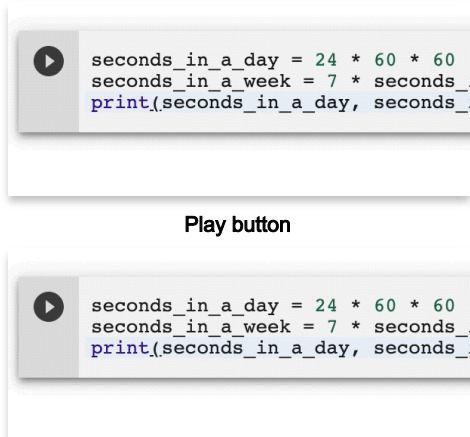
A Jupyter Notebook cell with a play button icon on the left. The cell contains two lines of Python code: `seconds_in_a_day = 24 * 60 * 60` and `print(seconds_in_a_day)`. Below the cell, a new cell is visible, containing the code `[] seconds_in_a_week = 7 * seconds_in_a_day` and `print(seconds_in_a_week)`. This illustrates the effect of pressing Shift+Enter, which runs the current cell and moves the cursor to the next cell below.

Shift + Enter



A Jupyter Notebook cell with a play button icon on the left. The cell contains two lines of Python code: `seconds_in_a_day = 24 * 60 * 60` and `print(seconds_in_a_day, seconds_in_a_week)`. Below the cell, a new empty code cell is inserted, illustrating the effect of pressing Alt+Enter, which runs the current cell and inserts a new cell below it.

Alt + Enter



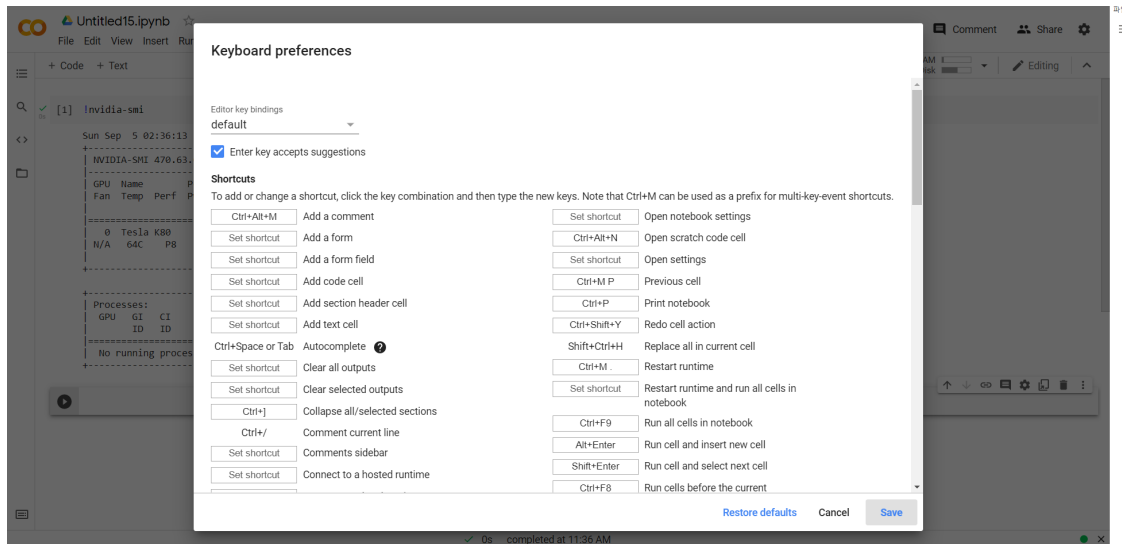
A Jupyter Notebook cell with a play button icon on the left. The cell contains two lines of Python code: `seconds_in_a_day = 24 * 60 * 60` and `print(seconds_in_a_day, seconds_in_a_week)`. Below the cell, another identical cell is shown, illustrating the effect of pressing Ctrl+Enter, which runs the current cell in place and inserts a new cell below it.

Play button

Ctrl + Enter

Working with Cells

- You can check list of keyboard shortcuts by typing `ctrl + m + h`





Working with Cells

- You can run any command on the system shell by prefixing it with

```
[1] !ls /
```

```
↳ bin      datalab  home    lib64  opt    run    swift      tmp    var
   boot     dev      lib     media  proc   sbin   sys        tools
   content  etc      lib32   mnt    root   srv    tensorflow-2.0.0-rc0  usr
```

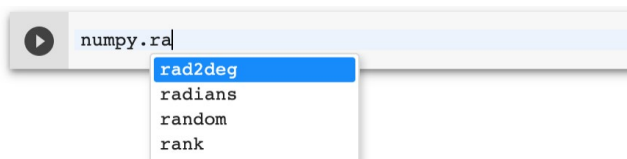
- You can install third-party libraries using package managers as needed

```
[ ] !apt install -y openjdk-11-jdk
    !pip install tqdm
```

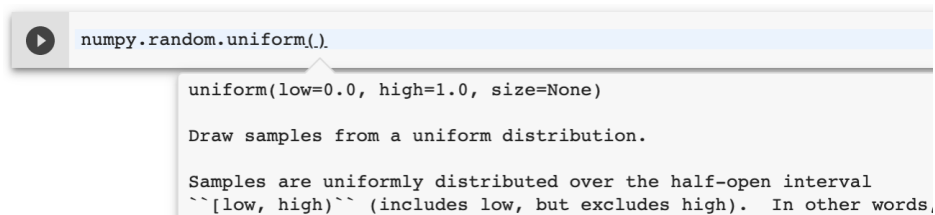
```
↳ Reading package lists... Done
   Building dependency tree
   Reading state information... Done
   The following additional packages will be installed:
     openjdk-11-jdk-headless
   Suggested packages:
     openjdk-11-demo openjdk-11-source visualvm
```

Working with Cells

- You can use tab completion to explore attributes of Python objects, as well as to quickly view documentation strings
 - As an example, if you press Tab after `numpy.ra`, you will see the list of available completions starting with `ra` within the `numpy` module



- You can see a popup of its documentation string



Working with Cells

- To open the documentation in a persistent pane at the bottom of your screen, add `?` after the object or method name and run the cell

The screenshot shows the Jupyter Notebook interface. At the top, there's a toolbar with icons for Comment, Share, and a user profile. Below that, a menu bar includes File, Edit, View, Insert, Runtime, Tools, and Help. A status bar shows RAM and Disk usage, and a mode selector set to 'Editing'. The main area contains two cells: a code cell with `[1] import numpy` and a help cell for `numpy.random.uniform?`. The help cell is expanded, showing the docstring for `uniform` in a pane at the bottom of the screen. The pane has tabs for History, Find/Replace, and Help (which is active). The docstring text is: `Docstring:`, `uniform(low=0.0, high=1.0, size=None)`, and `Draw samples from a uniform distribution.`

```
[1] import numpy
```

```
numpy.random.uniform?
```

Docstring:
`uniform(low=0.0, high=1.0, size=None)`
Draw samples from a uniform distribution.



Google drive in Colab

- You can mount your google drive in Colab
 - If you do NOT use the google drive, you may lose your all data and results when Colab's session expired.

Access to your google drive in Colab

```
from google.colab import drive  
  
drive.mount('/gdrive')
```

```
# check what's in the mounted gdrive using Colab  
import os  
  
gdrive_root = '/gdrive/My Drive'  
print('In gdrive :', os.listdir(gdrive_root))  
  
notebook_dir = os.path.join(gdrive_root, 'ColabNotebooks')  
print('In Colab Notebooks :', os.listdir(notebook_dir))
```

Fetching Notebooks from GitHub

- You can fetch notebooks from GitHub repositories by searching them
 - In this lecture, we will upload all the notebooks to repositories by keai - kaist

The screenshot shows the GitHub search page. At the top, there is a navigation bar with tabs: '예', '최근 사용', 'Google Drive', 'GitHub' (highlighted with a green box and labeled '1'), and '업로드'. Below the navigation bar, there is a search bar with the text 'GitHub URL을 입력하거나 조직 또는 사용자로 검색하세요.' and a checkbox for '비공개 저장소 포함'. The search bar contains the text 'keai-kaist' (labeled '2'). Below the search bar, there is a dropdown menu showing search results (labeled '3'). The first result is 'keai-kaist/CS470-Spring-2022-' with a blue highlight on 'keai-kaist/CS470-Spring-2022-'. Below the dropdown menu, there is a list of repositories. The first repository is 'Lab1/Apr 26/Class1-1.ipynb' and the second is 'Lab1/Apr 26/Class1-2.ipynb'. To the right of each repository name, there is a checkbox (labeled '4') and a document icon.