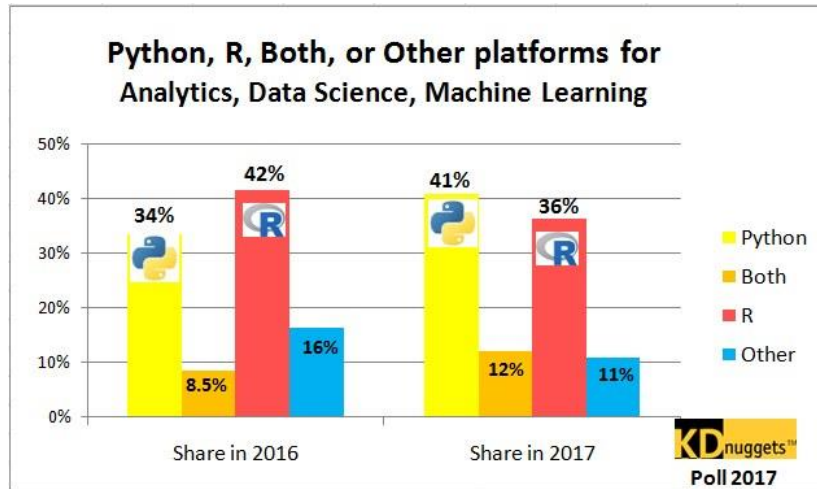




1-1

Google Colab 사용법

Programming language for Machine Learning



출처 - <http://artificialintelligencemania.com/2018/07/02/the-best-programming-language-for-machine-learning/>



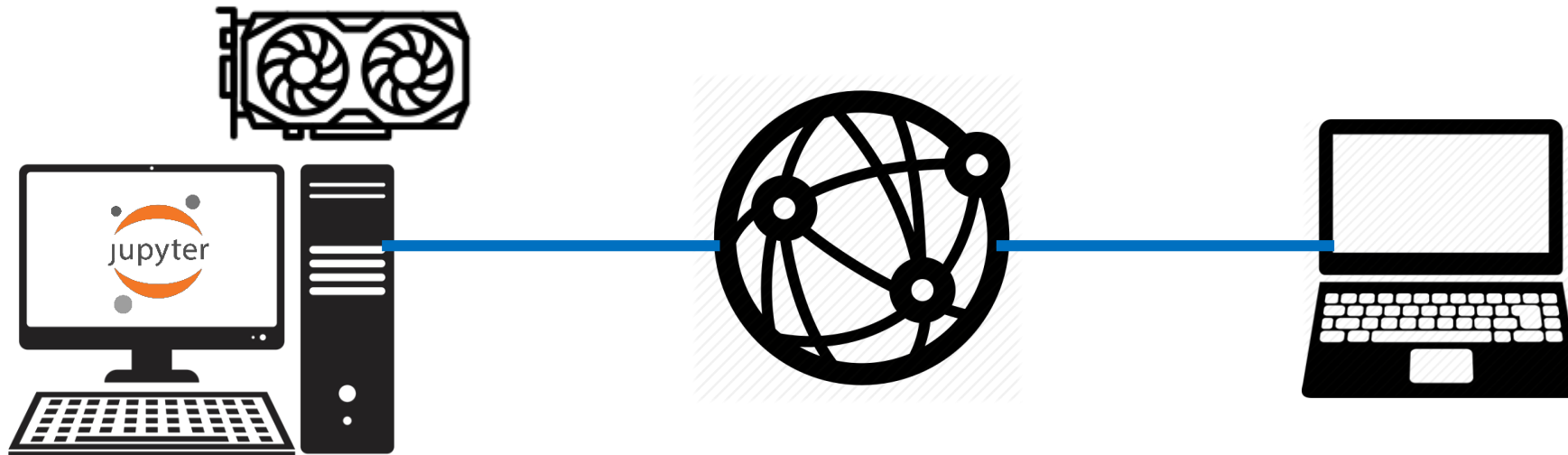
■ Python

- ML 연구 분야에 있어서 대체하기 어려운 프로그래밍 언어
- Tensorflow, Pytorch 등의 딥러닝 프레임워크
- Numpy, Jupyter Notebook, Matplotlib, Pandas, ...



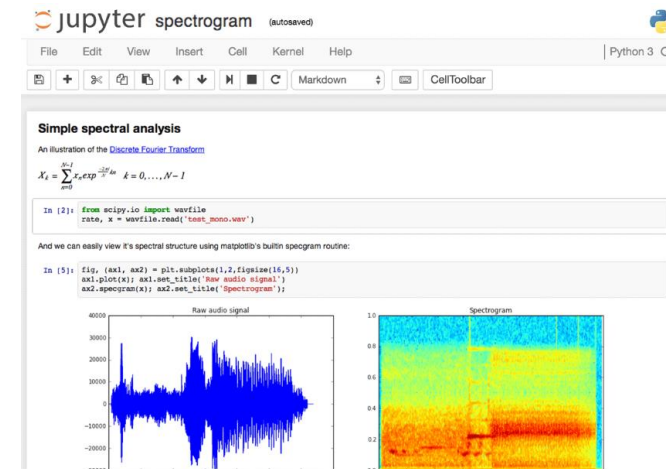
Jupyter Notebook?

- 웹 브라우저에서 파이썬 코드를 작성하고 실행해 볼 수 있는 개발도구
 - 원격 코딩 가능
 - 코드 블록 단위로 실행 / 디버깅
 - Text block을 이용한 문서화
 - Figure plotting 등 GUI



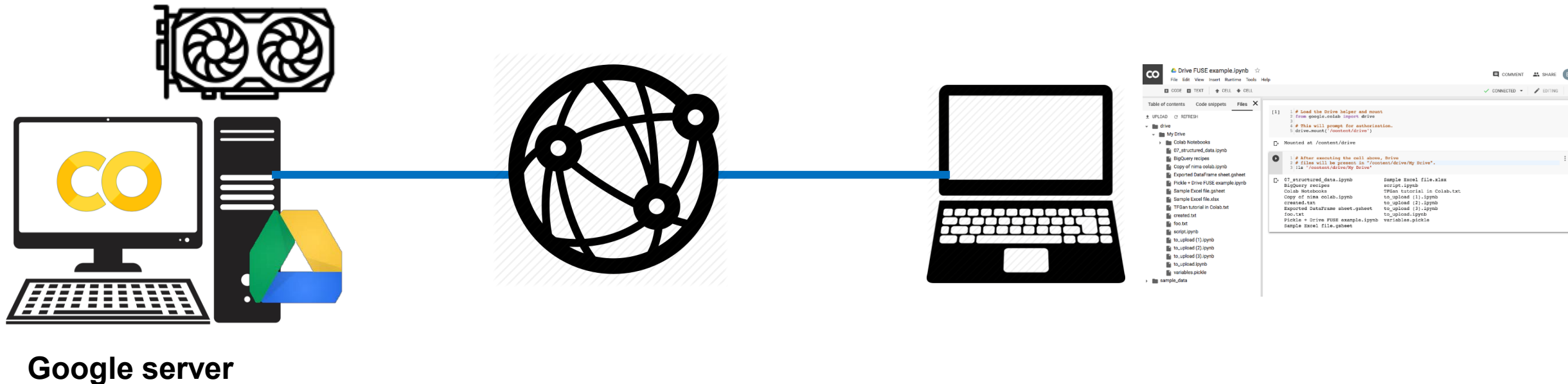
147.46.123.123

147.46.123.123:8888



Google Colab?

- Google Colaboratory = Google Drive + Jupyter Notebook
 - 구글 계정 전용의 가상 머신 지원 – **GPU 포함**
 - Google drive 문서와 같이 링크만으로 접근 / 협업 가능
 - 코드 실행 시 딜레이 존재



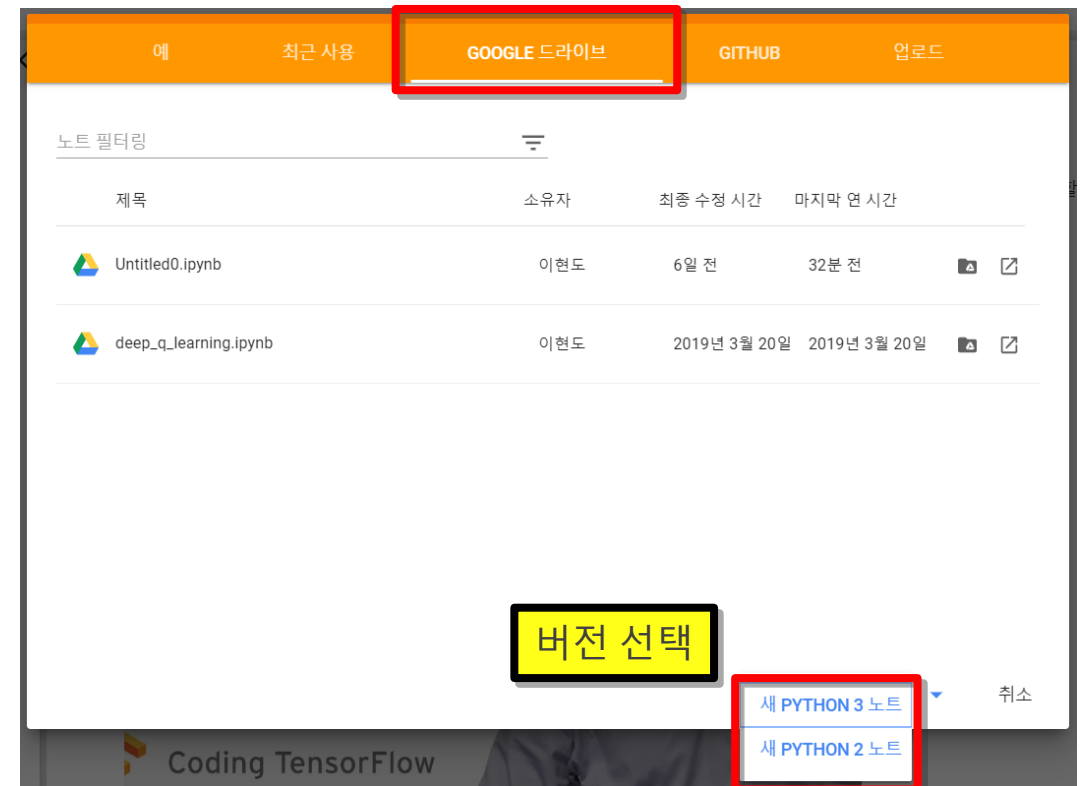
Google Colab - 사용법

- 개인 구글 계정 필요
- Colab과 Jupyter Notebook 사용 방법은 유사한 부분이 많음
 - 실습 수업에서는 Colab 위주로 설명
 - GPU가 내장된 서버를 사용할 수 있을 시 로컬에서 작업을 권장

Google Colab - 사용법

■ 파일 생성/접근 방법

- 개인 구글 계정으로 접속
- <https://colab.research.google.com> 접속
- GOOGLE 드라이브 탭 이동
- 새 PYTHON 노트 선택
 - 실습은 python3로 진행할 예정



Google Colab - 사용법

■ 파일 이름 변경



■ Code cell, Text cell

- .ipynb 파일은 code cell과 text cell로 구성
- 각 셀 하단에 마우스를 대거나, 화면 좌상단 버튼으로 셀 추가 가능
- 셀 선택(마우스) 후 셀 우상단 삭제버튼으로 셀 삭제 가능

Google Colab - 사용법

■ Code cell

- 일반적인 파이썬 코딩 방식과 동일
- 각 셀은 한번에 실행할 단위를 뜻함
- 실행 이후에도 메모리는 유지되어 다른 셀 실행 시 영향을 줌
 - 런타임 다시 시작 시 초기화

- 상단 메뉴의 런타임
 - 실행 중인 셀 중단
 - 런타임 다시 시작

The screenshot shows a code cell with two blocks of Python code. The first block is labeled with a yellow box '실행 번호' (Execution Number) and '[1]'. The code is: `# Code Cell!
a = 1
b = 2
print(a+b)`. Below it, a comment says '# Ctrl+Enter 로 해당 코드 셀 실행'. The output is '3'. The second block is labeled with a yellow box '실행 (Ctrl + Enter)' and contains: `# 각 셀은 한번에 실행할 단위를 뜻함
실행 이후에도 메모리는 그대로 유지되어 다른 셀의 실행에 영향을 줌
a += 3
b -= 1
print(a+b)`. The output is '5'. A yellow box '실행 결과' (Execution Result) points to the output '5'.

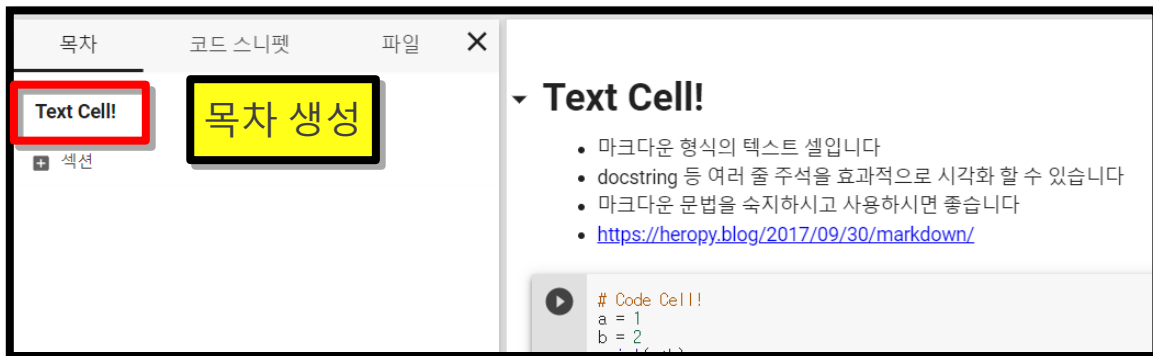
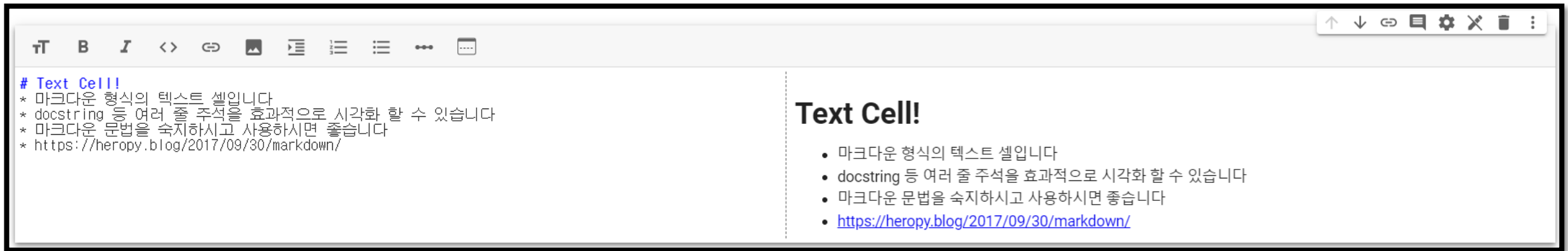
The screenshot shows the 'Run' menu (labeled '런타임' in a yellow box) with various options. The 'Interrupt current cell' (실행 중단) option is highlighted with a red box, with the keyboard shortcut 'Ctrl+M' shown next to it. Below it, the 'Restart runtime' (런타임 다시 시작...) option is also highlighted with a red box, with the keyboard shortcut 'Ctrl+M' shown next to it. Other options include 'Run all cells' (모두 실행), 'Run previous cell' (이전 셀 실행), 'Run selected cells' (선택항목 실행), 'Run from here' (이후 셀 실행), 'Restart all cells' (다시 시작 및 모두 실행...), and 'Reset all runtimes' (모든 런타임 재설정...).

실행 중단
런타임 재시작

Google Colab - 사용법

■ Text cell

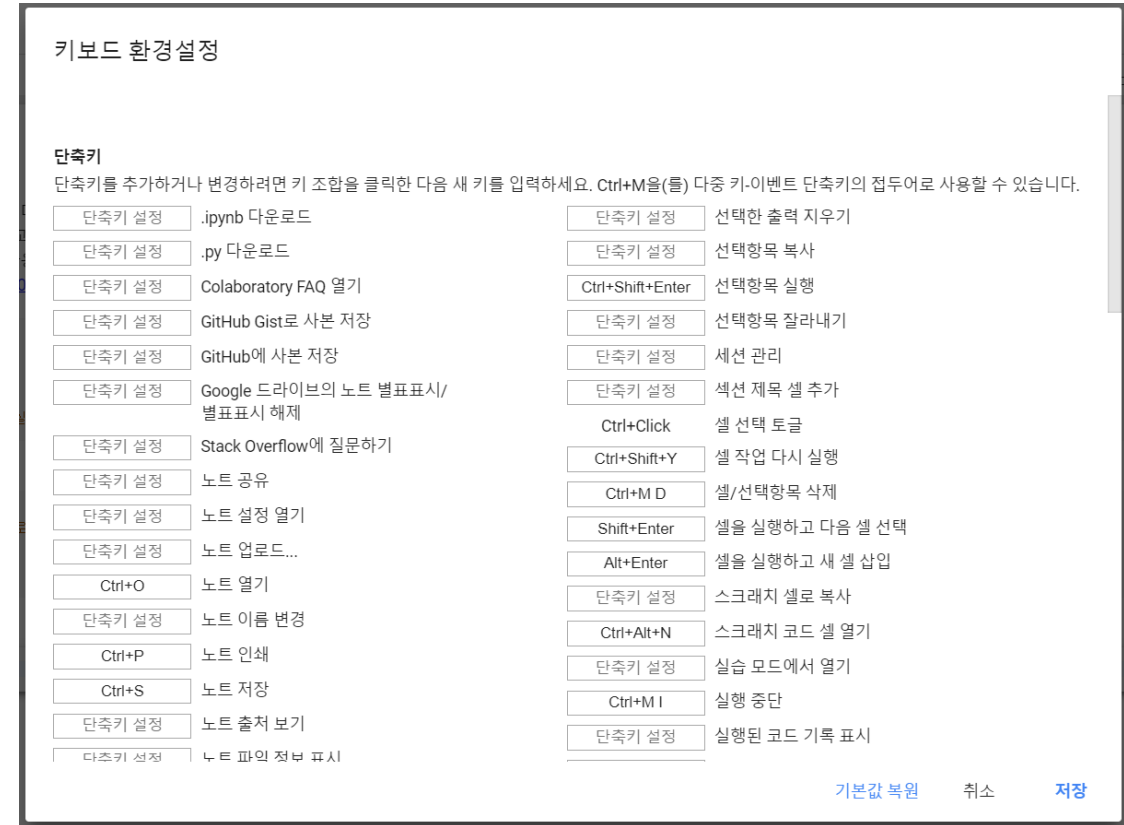
- 여러 줄 주석의 효과적인 시각화
- 마크다운(Markdown) 문법
- 자동 목차 생성



Google Colab - 사용법

■ 단축키

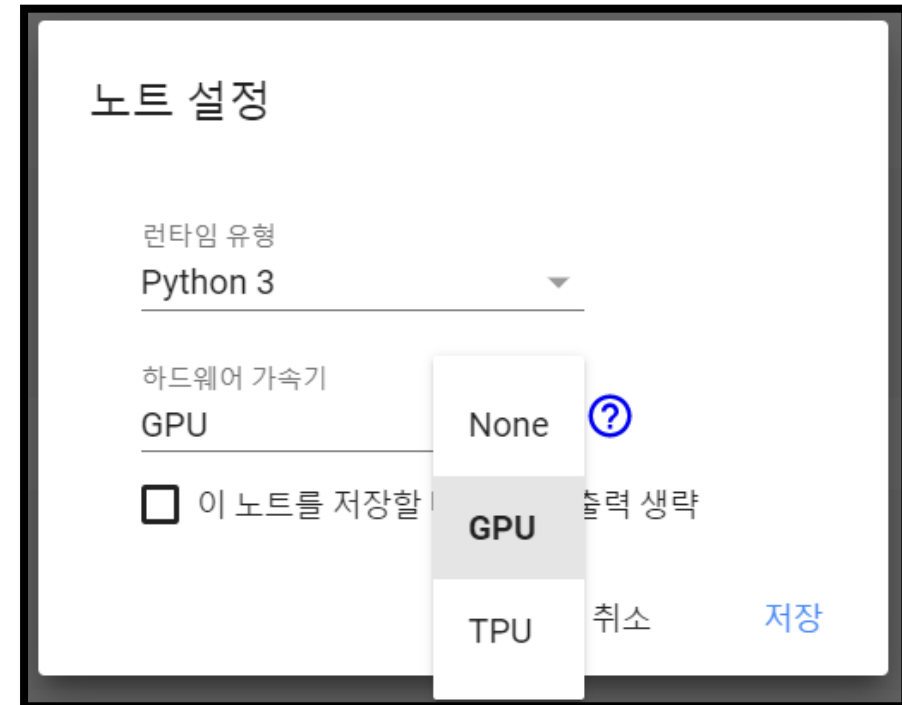
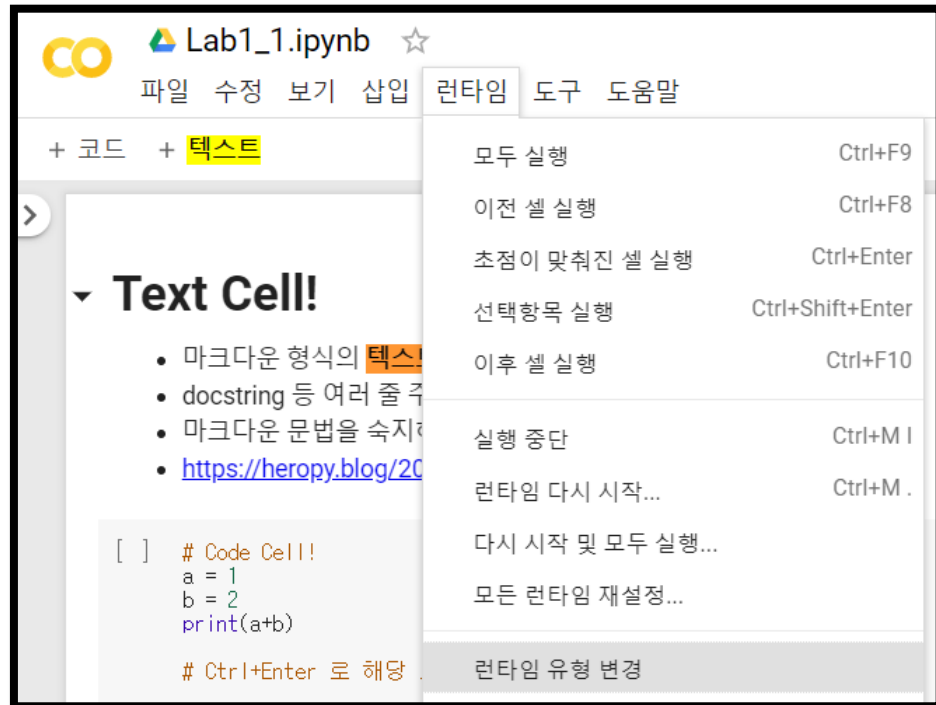
- 대부분의 작업은 단축키로 실행 가능
- 단축키 설정 가능
- 단축키 설정화면 – Ctrl+M H
- 유용한 단축키
 - 코드 셀 생성 – Ctrl+M A(B)
 - 코드 셀 실행 – Ctrl+Enter
 - 셀 삭제 – Ctrl+M D
 - 실행중인 셀 중단 – Ctrl+M I
 - 런타임 다시 시작 – Ctrl+M .
 - 코드(텍스트) 셀로 변환 – Ctrl+M Y(M)
 - 마지막 셀 작업 실행취소 – Ctrl+Shift+Z



Google Colab - 사용법

■ GPU 설정

- 런타임 -> 런타임 유형 변경 -> 하드웨어 가속기를 GPU로 변경
- 유의사항 – GPU는 최대 12시간 실행을 지원
 - 12시간 실행 이후에는 런타임 재시작으로 VM을 교체해야 함



Google Colab - 사용법

■ 명령어 실행하기

- !코드 셀에 를 붙이고 터미널 명령어를 입력하여 실행하면 터미널에서 실행하는 것과 같은 결과가 출력됨
- 예외로 cd 명령어는 %cd /your/desired/path

```
!cat /etc/issue.net # OS
!cat /proc/cpuinfo # CPU
!cat /proc/meminfo # Memory
!df -h # Disk
!nvidia-smi # GPU
```

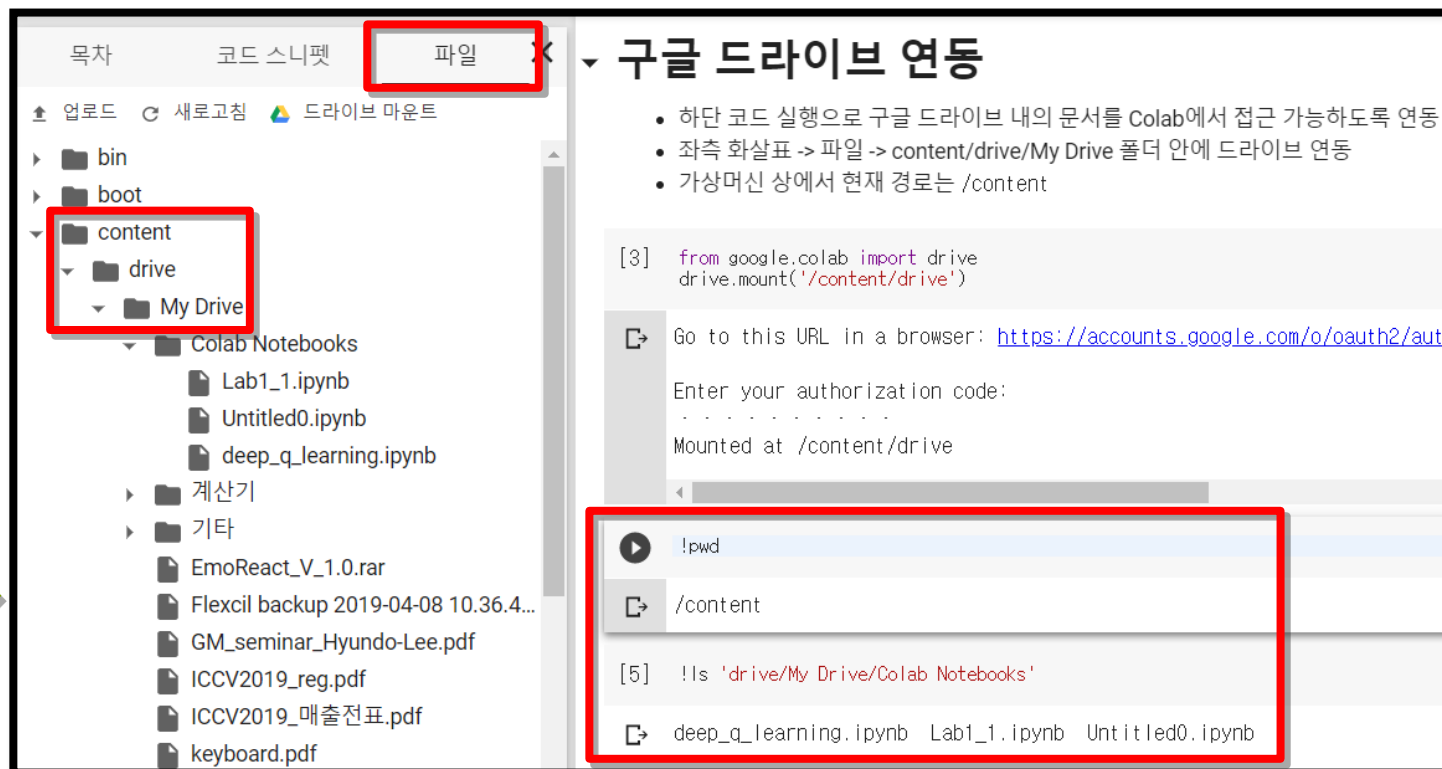
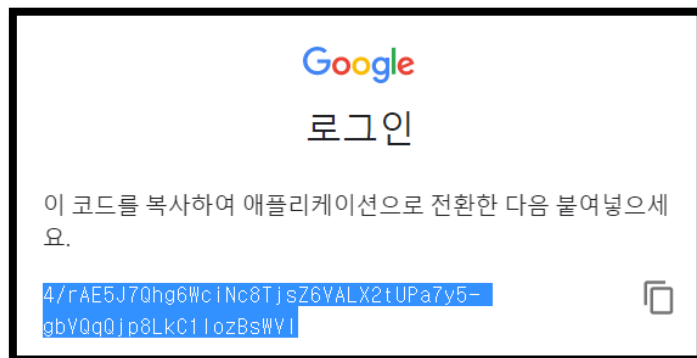
Mon Sep 16 07:49:42 2019

+-----+-----+-----+-----+									
NVIDIA-SMI		430.40		Driver Version: 418.67		CUDA Version: 10.1			
+-----+-----+-----+-----+									
GPU	Name	Persistence-M	Bus-Id	Disp.A	Volatile Uncorr. ECC				
Fan	Temp	Perf	Pwr:Usage/Cap	Memory-Usage	GPU-Util	Compute M.			
+-----+-----+-----+-----+									
0	Tesla K80	Off	00000000:00:04.0 Off			0			
N/A	52C	P8	32W / 149W	0MiB / 11441MiB	0%	Default			
+-----+-----+-----+-----+									
+-----+-----+-----+-----+									
Processes:						GPU Memory			
GPU	PID	Type	Process name			Usage			
+-----+-----+-----+-----+									
No running processes found									
+-----+-----+-----+-----+									

Google Colab - 사용법

■ 구글 드라이브 연동

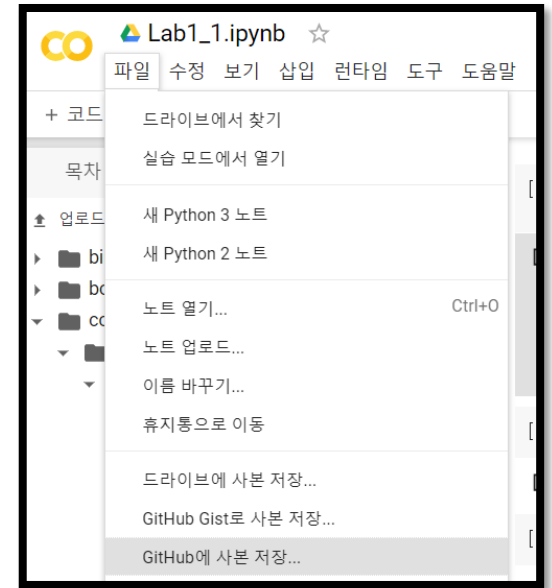
- 간단한 인증 절차 이후 구글 드라이브의 파일을 Colab에서 접근 가능



Google Colab - 사용법

■ Github 연동

- 단일 .ipynb 파일을 clone 하는 방법
 - `https://github.com/~~~` 부분을
`https://colab.research.google.com/github/~~~` 로 교체
 - 파일 -> 드라이브에 사본 저장
- 전체 repository cloning
 - `!git clone project.git`
- github repository에 파일을 올리는 방법
 - 파일 - Github에 사본 저장 선택
 - 저장소, 브랜치, 경로 지정



GitHub으로 복사

저장소: 브랜치:

파일 경로
Lab1_1.ipynb

변경사항 설명 메시지
Colaboratory를 통해 생성됨

☒ Colaboratory 링크 추가

취소 확인

Google Colab - tips

■ 장시간 사용할 때의 issue들

- The '**maximum lifetime**' of a running notebook is **12 hours** (browser open)
- An '**Idle**' notebook instance cuts-off after **90 minutes**
- You can have a maximum of **2 notebooks** running concurrently
- If you close the notebook window and open it while the instance is still running, the cell outputs and variables will still persist. However if the notebook instance has been recycled, your cell outputs and variables will no longer be available.

- GPU 사용/미사용 관계없이 최대 12시간
- 아무것도 안 하는 idle 상태 돌입 이후 90분에 런타임 자동으로 shutdown
 - Shutdown 전에 창을 끄고 다시 켜도 런타임은 유지됨
- 동시에 최대 2개의 notebook 가동 가능

Google Colab - tips

- Jupyter로 코딩하고 싶은데 Colab GPU만 쓰고 싶은 경우
 - 로컬에서 작업 후 drive에 올리고 Colab으로 실행
 - 학습된 모델 weight만 다운로드 후 로컬에서 분석
- Jupyter도 쓰기 싫은 경우
 - Drive에 올리고 실행해야 할 파일(ex: train.py)를 Colab에서 커맨드로 실행(!train.py)
- 그냥 Colab 쓰는게 싫은 경우
 - 학기마다 GPU 서버 사용 신청을 할 수 있음(서울대학교 구성원)

References

- <https://zzsza.github.io/data/2018/08/30/google-colab/>
- <https://stackoverflow.com/questions/55050988/can-i-run-a-google-colab-free-edition-script-and-then-shutdown-my-computer>