

생물공학 x 기계학습 워크샵 필수 사전 준비사항

안녕하세요, 12월 19일 생물공학x기계학습 워크샵 실습 교육 관련 사전 준비사항을 안내해 드립니다. 개인 컴퓨터를 이용한 실습을 효율적으로 원활히 진행하기 위해서는 **사전에 실습 환경을 철저히 준비**해 오시는 것이 중요합니다.

기본 실습 환경 설정은 Docker 로서 다음 매뉴얼을 따라 수행해 보시고 하드웨어 문제등으로 부득이 설치가 어려울 경우 colab 환경에서 실습을 진행하겠습니다. 매뉴얼 후반부에 소개되어 있으며 **colab 환경은 인터넷만 연결되어 있으면 사용 가능하나 12시간 후 초기화 되며 실행이 느릴 수 있으니** 참고하시길 바랍니다.

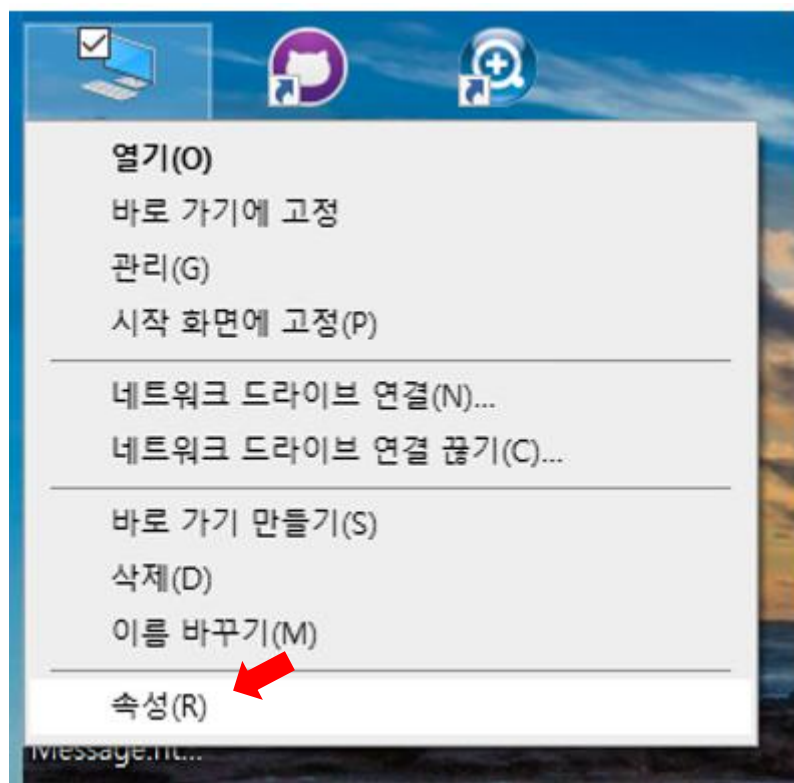
참고로 모든 실습은 Jupyter notebook 으로 진행될 예정이므로 인터넷을 참고하시어 사전에 사용법을 익혀오시면 진행에 도움이 되겠습니다.

실습을 위해서 가능한 최신 윈도우 노트북 컴퓨터를 사용 하시기를 추천 드립니다

확인 방법

내컴퓨터 > 오른쪽 마우스버튼 > 속성

Windows 10 64bit Pro 이상



Windows 버전

Windows 10 Pro

© 2019 Microsoft Corporation. All rights reserved.

시스템

| | |
|---------------|---|
| 제조업체: | Lenovo |
| 모델: | ThinkPad T470 Signature Edition |
| 프로세서: | Intel(R) Core(TM) i7-7600U CPU @ 2.80GHz 2.90 GHz |
| 설치된 메모리(RAM): | 32.0GB (11.9GB 사용 가능) |
| 시스템 종류: | 64비트 운영 체제, x64 기반 프로세서 |
| 펜 및 터치: | 10개의 터치 포인트를 사용할 수 있는 터치 지원 |

최종 사전 준비 완료 상태

- 도커 설치하신 분들

1) docker 설치 완료

2) docker 이미지 bioengml:v1 다운로드 완료

다음 명령어로 확인

```
PS C:\bioengml> docker images
```

| REPOSITORY | TAG | IMAGE ID | CREATED | SIZE |
|------------------|-----|--------------|------------|--------|
| haseong/bioengml | v1 | e6fc522deb48 | 3 days ago | 5.61GB |

- 도커 설치 불가능하신 분들 colab 이용

1) 구글 계정 준비

2) 당일 수업 전 구글드라이브 연결 및 다음 패키지 설치

```
!pip install biopython
```

목 차

1. Docker 환경 설정 (추천, 윈도우, MacOS)

1. git 설치
2. docker 설치를 위한 필수 사항 확인
3. docker 설치
4. docker 이미지 다운로드
5. Chrome 브라우저 설치
6. 컨테이너 실행 및 예습

2. colab 환경 설정

1. gmail 계정 만들기
2. chrome 브라우저 설치
3. colab 접속
4. 구글 드라이브 연결
5. 수업 내용 다운로드

Docker 환경 설정

docker는 리눅스 기반 가상화 플랫폼으로 빠르고 효율적인 컴퓨팅 환경 구축이 가능하므로 최근 주목을 받는 기술 중 하나입니다. 본 워크샵에서 수행하는 기계학습 실습 뿐만 아니라 딥러닝 등 다양한 데이터 분석 환경을 구축하는 과정은 데이터 분석 자체 보다 더 복잡하고 까다로울 수 있습니다. 도커의 사용법을 학습해 두면 앞으로 더 복잡해지는 컴퓨팅 환경에 대한 부담을 줄이고 본연의 데이터 분석과 해석에 더 충실할 수 있을 것입니다.

Windows

1. git 설치

- <https://git-scm.com/downloads> 에서 Windows 클릭
- 설치파일 실행
- 모든 옵션을 default 로 하고 Next 클릭
- 설치 후 임의 디렉토리에서 shift + 마우스 오른쪽버튼 클릭
- 메뉴 중 “여기에 Powershell 창 열기” 클릭
- PowerShell에서 git 입력 후 실행 확인

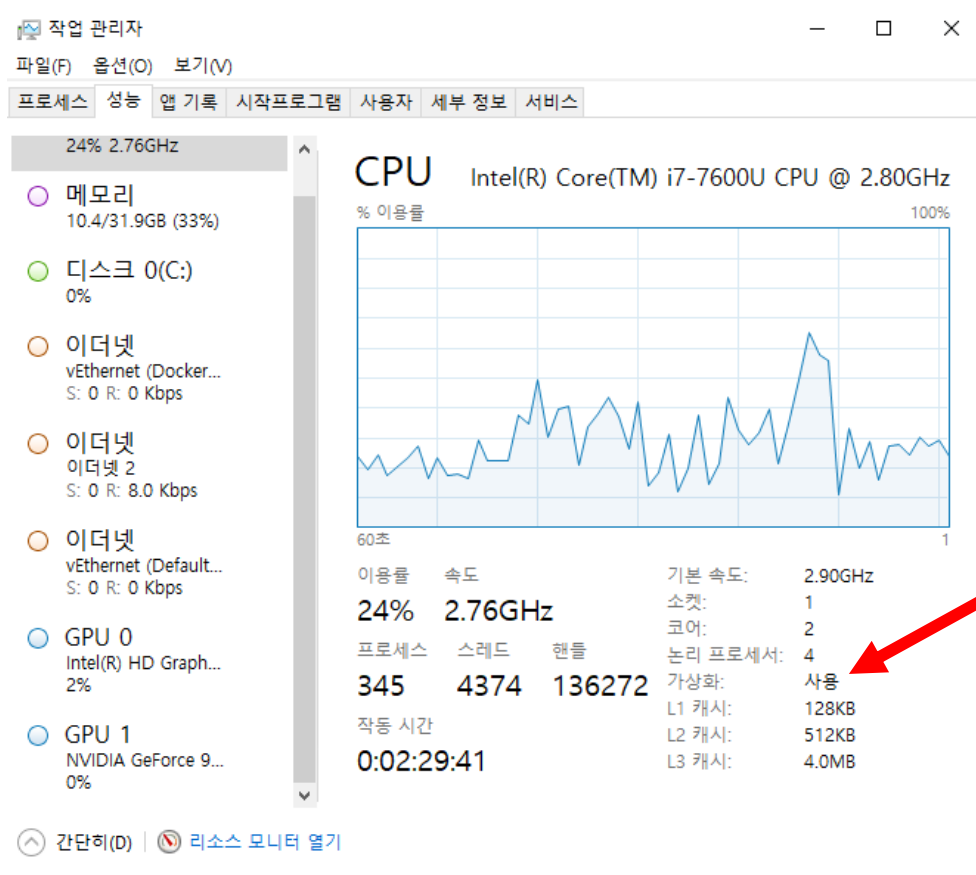
2. Docker 설치 전 확인 사항

- 다음 두 가지 사항 필수
 - Windows 10 64-bit Pro 이상 (Home edition 사용 불가)
 - Hyper-V 사용 가능
- Hyper-V 사용을 위해서는 다음 조건 필요
 - 64bit processor
 - 4GB 이상 RAM
 - BIOS-level 하드웨어 가상화 지원 (CPU virtualization)

BIOS level 하드웨어 가상화부터 다음 페이지 설명

<https://docs.docker.com/docker-for-windows/install/> 참고

2.1 하드웨어 가상화 지원 확인

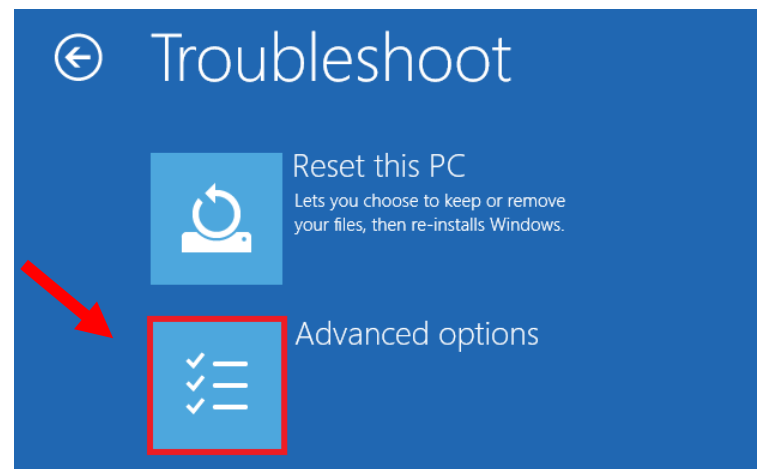
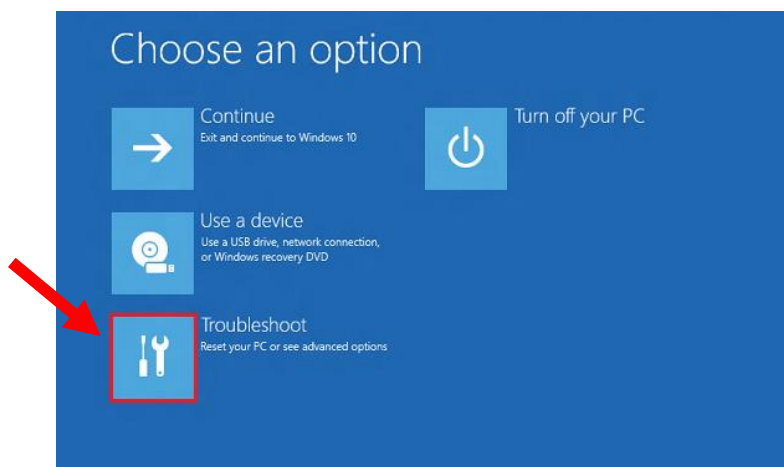


“사용”으로 되어
있어야 함

“사용”이 아닐 경우 BIOS 설정 필요, 다음 페이지

2.2 BIOS에서 가상화 설정 방법

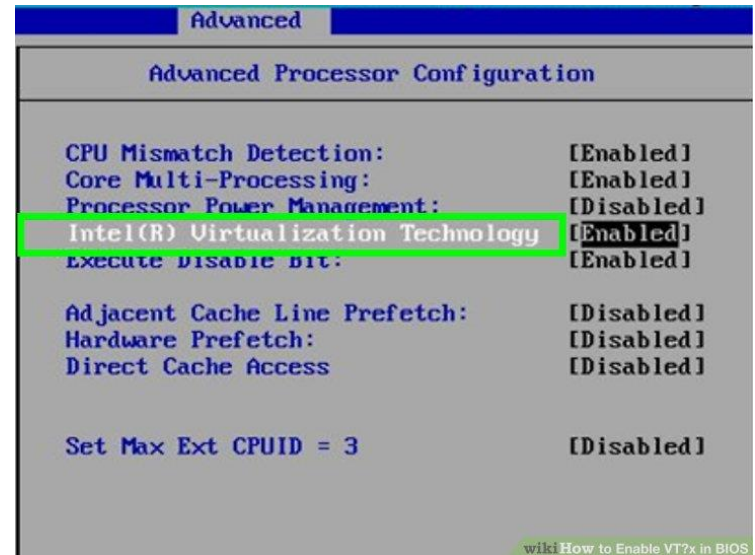
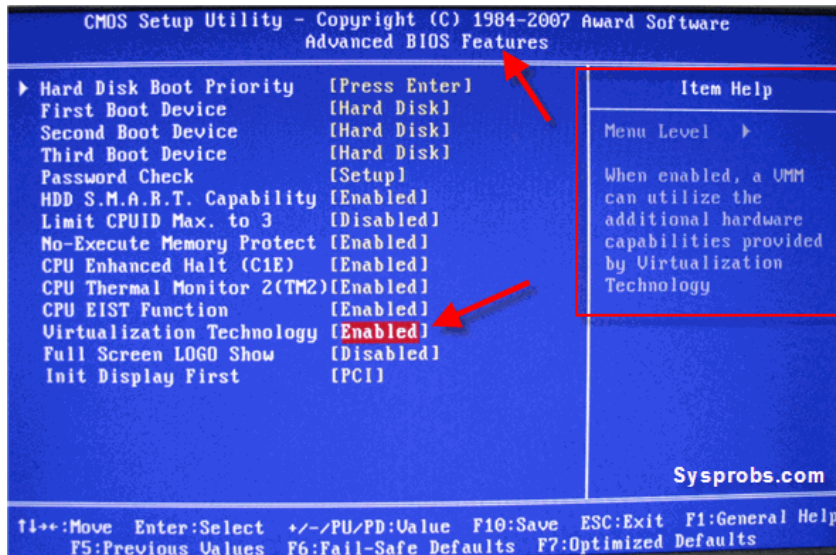
- 윈도우 시작버튼 > 설정 > 업데이트 및 보안 > 복구 > 지금 다시 시작
- 위 과정 진행하면 다음 화면 나옴 (작업 중인 내용 종료 주의)



- 문제 해결 > 고급 옵션 > UEFI 펌웨어 설정 선택 후 “다시시작” 버튼 클릭
- 컴퓨터가 다시 시작하며 자동으로 BIOS로 진입

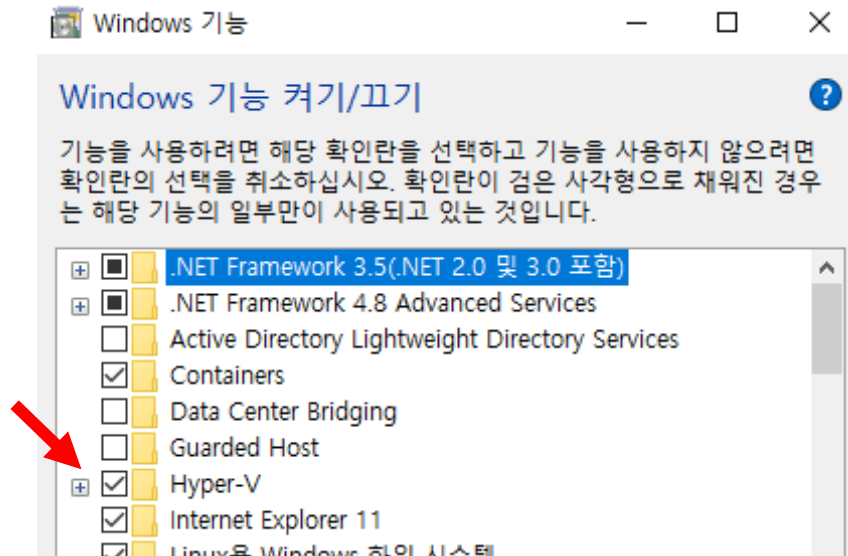
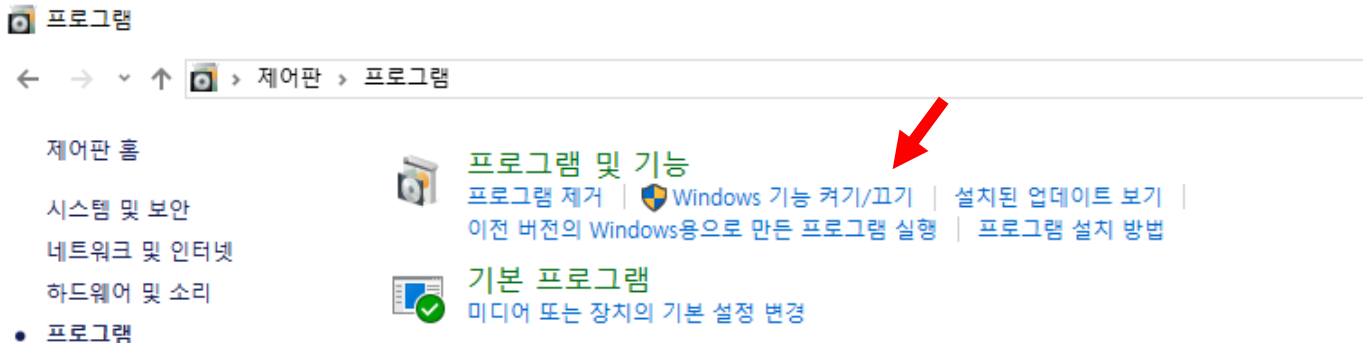
2.2 BIOS에서 가상화 설정 방법

- BIOS에서 CPU 가상화 설정법은 컴퓨터(메인보드)마다 다를 수 있음.
아래 그림들은 일부 예시 참고 그림
- Advanced BIOS features > CPU 관련 메뉴 > Virtualization Technology 또는 SVM mode 항목 > Enable 시킨 후 저장 > 재부팅
- 구글에서 “bios virtualization enable” 검색 참고



2.3 MS windows 가상화 확인

- 윈도우 > 제어판 > 프로그램 > windows 기능 켜기/끄기

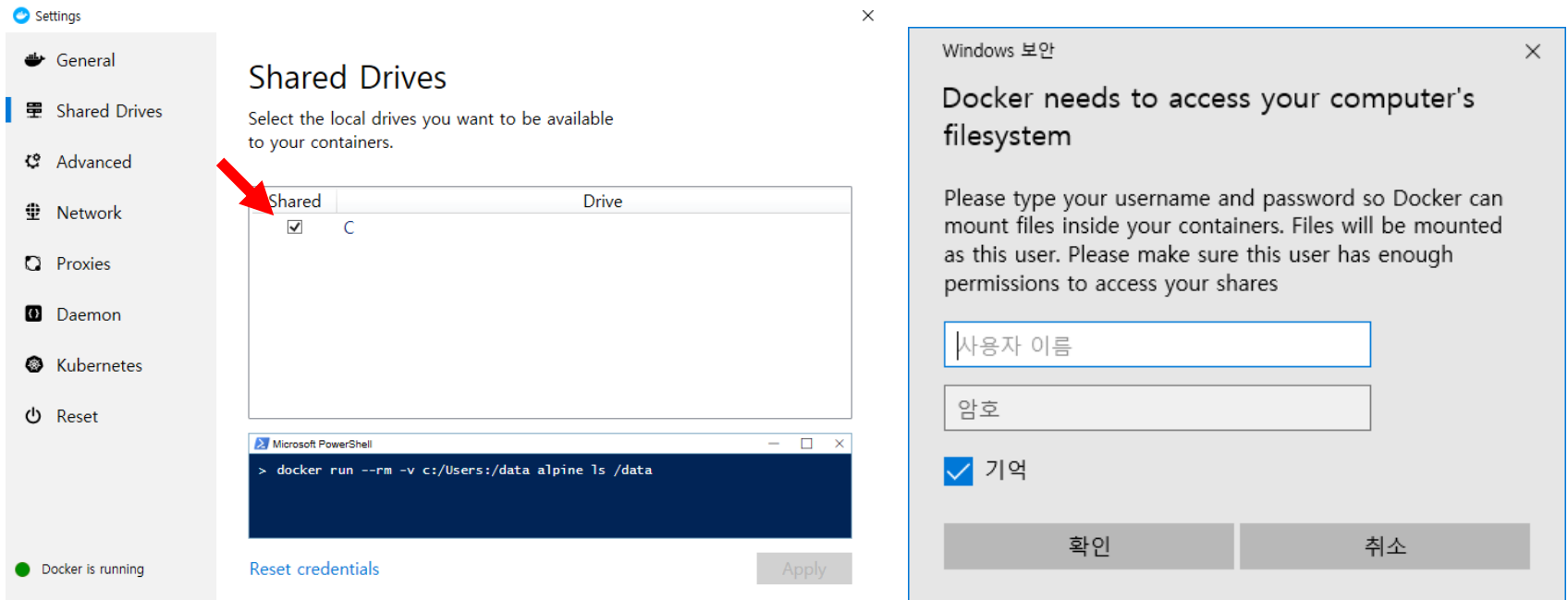


3.1 Docker 설치

- Docker 윈도우즈 설치 참고 <https://docs.docker.com/docker-for-windows/>
- (계정이 없을 경우) Docker Hub (<https://hub.docker.com/>) Get Started 클릭 후 사용자 등록
- (계정이 있을 경우) <https://www.docker.com> > Products > Desktop > Download Desktop for Mac and Windows 클릭 > 새 창 열리며 Sign in
- 빠른 시작 화면에서 설치파일 (Download Desktop for Windows) 다운로드
- 설치파일 실행 (설치 후 튜토리얼 계속 실습 추천)

3.2 디렉토리 공유 설정

- 컨테이너 저장소는 휘발성으로 컨테이너 종료와 동시에 사라짐. 따라서 컨테이너 저장소와 하드드라이브 공유 필요
- 윈도우 하단 도커 아이콘 > 마우스 오른쪽버튼 > Settings > Shared Drives
- 아래와 같이 타겟 드라이브 공유
- 윈도우 사용자 계정 입력 (비밀번호 필수, 없을 경우 설정 필요)

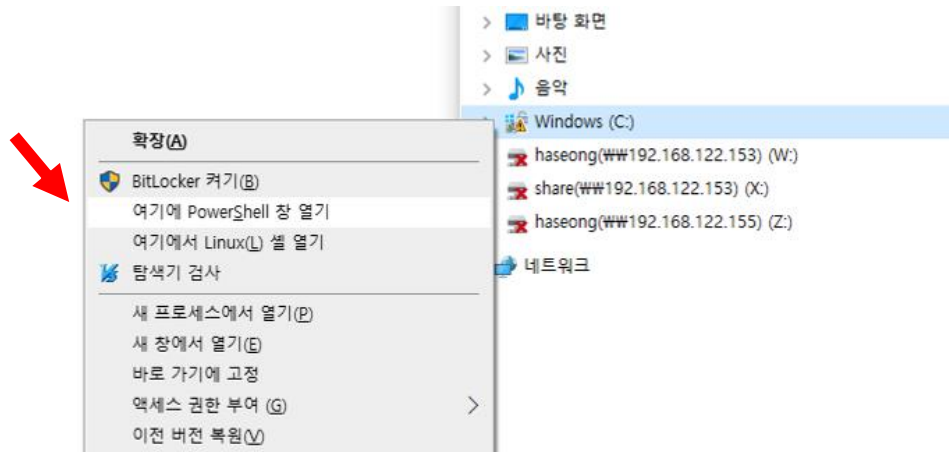


3.3 방화벽 이슈

- docker는 임의로 윈도우의 host 파일을 변경하여 알집이나 V3에 의해서 악성코드로 감지당할 수 있습니다
- V3의 경우 파일 변경을 감지하고 원래 파일로 되돌릴지 물어볼 경우 “아니오” 를 선택합니다
- 알집의 경우 환경설정에서 호스트파일 보호 옵션을 off 하거나 파일 예외 처리를 하면 되겠습니다
- 다른 보안 이슈들이 있을 경우 검색이나 문의를 통해 조치 가능합니다

4. 도커 이미지 다운로드

- 브라우저의 c:\W 드라이브에서 Shift + 마우스오른쪽버튼 클릭
PowerShell 실행 (c:\Wbioengml\W 를 작업 디렉토리로 가정)



- 다음 명령으로 필요 파일들 복사 및 이미지 다운로드 (약 5.3GB)

```
PS C:\> git clone https://github.com/greendaygh/bioengml.git
PS C:\> cd bioengml
PS C:\bioengml> docker pull haseong/bioengml:v1
PS C:\bioengml> docker images
```


4. (참고) 도커 이미지 직접 생성

- 이미지를 다운로드 하지 않고 Dockerfile 이용하여 직접 생성 가능
- Dockerfile 작성 방법은 인터넷 또는 다음 레퍼런스 페이지 참고

<https://docs.docker.com/engine/reference/builder/>

```
PS C:\> cd bioengml
```

```
PS C:\bioengml> dir
```

| Mode | | LastWriteTime | Length | Name |
|--------|--|--------------------|--------|----------------|
| d----- | | 2019-11-30 오후 6:08 | | datasets |
| -a---- | | 2019-11-30 오후 6:14 | 1644 | Dockerfile |
| -a---- | | 2019-11-30 오후 6:55 | 205 | docker_run.bat |
| -a---- | | 2019-11-30 오후 6:09 | 154 | README.md |

```
PS C:\bioeng-m1> docker build -t haseong/bioengml:v1 .
```

```
PS C:\bioeng-m1> docker images
```

5. Chrome 브라우저 설치 (없을경우)

- 본 워크샵에서는 Jupyter notebook으로 주요 실습이 진행될 예정입니다. Jupyter notebook의 경우 Internet explorer에 대해서 지원이 잘 안되는 문제로 Chrome 브라우저 설치가 필요합니다
- 설치는
https://support.google.com/chrome/answer/95346?hl=ko&ref_topic=7439538# 이곳을 참고 하세요

6. 컨테이너 실행 및 예습

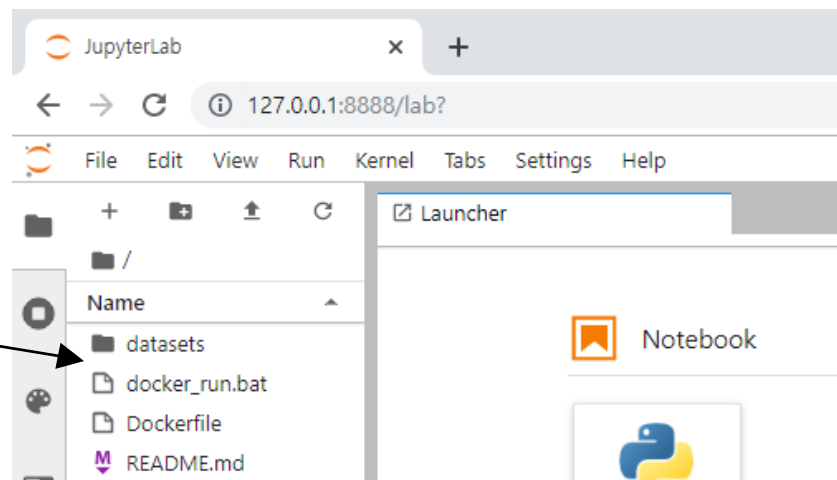
- github에서 다운로드받은 파일 중 docker_run.bat 배치파일 실행

```
PS C:\bioengml> .\docker_run.bat
```

```
C:\bioengml>docker run --rm -it -d -p 8888:8888 --name bioengml -v  
c:\bioengml:/home/bioengml haseong/bioengml:v1 jupyter lab --no-  
browser --allow-root --ip=0.0.0.0 --notebook-dir=/home/bioengml --  
NotebookApp.token=
```

- 크롬 브라우저를 열고 주소창에 127.0.0.1:8888 입력

작업 디렉토리
c:\WbioengmlW



6. (참고) 컨테이너 접속 외

- 컨테이너 실행 확인

```
PS C:\bioengml> docker ps
```

| CONTAINER ID | IMAGE | COMMAND | CREATED | ... | NAMES |
|--------------|---------------------|--------------------------|----------------|-----|----------|
| 892ea16386b9 | haseong/bioengml:v1 | "jupyter lab --no-br..." | 29 minutes ago | ... | bioengml |

- 컨테이너 접속

```
PS C:\bioengml> docker exec -it bioengml /bin/bash
```

```
root@892ea16386b9:/home/bioengml# ll
```

```
total 17
```

```
drwxrwxrwx 2 root root 8192 11월 30 12:43 .
drwxr-xr-x 1 root root 4096 11월 30 09:34 ..
drwxrwxrwx 2 root root    0 11월 30 12:43 .git
-rwxr-xr-x 1 root root 1644 11월 30 09:14 Dockerfile
-rwxr-xr-x 1 root root  154 11월 30 09:09 README.md
drwxrwxrwx 2 root root    0 11월 30 09:08 datasets
-rwxr-xr-x 1 root root  205 11월 30 09:55 docker_run.bat
```

6. (참고) 컨테이너 접속 외

- 컨테이너 접속 후 conda 환경 활성화

```
root@892ea16386b9:/home/bioengml# conda env list
# conda environments:
#
base                  *  /usr/local

root@892ea16386b9:/home/bioengml# source activate
(base) root@892ea16386b9:/home/bioengml#
```

- 컨테이너 나가기: Ctrl + D
- 컨테이너 종료

```
PS C:\bioengml> docker stop bioengml
```

- 참고로 컨테이너 접속 후 apt-get, conda, pip 등으로 설치한 패키지는 컨테이너가 종료되면 사라지므로 Dockerfile로 이미지 재생성 하거나 docker commit 이용

MacOS

1. git 설치

- <https://git-scm.com/book/ko/v2/>에서 1.5 git 설치 클릭
- 터미널에서 git 입력 후 실행 확인

Mac에 설치

Mac에 Git을 설치하는 방법 중에는 Xcode Command Line Tools를 설치하는 방법이 가장 쉽다. Mavericks(10.9)부터는 Terminal에 단지 처음으로 'git'을 실행하는 것으로 설치가 시작된다. 'git'이 설치돼 있지 않으면 설치하라고 안내해준다.

```
$ git --version
```

만약 Git이 시스템에 설치되어있지 않은 경우, 설치할 수 있도록 안내 메시지가 뜰 것이다.

좀 더 최신 버전이 필요하면 바이너리 인스톨러로 설치할 수 있다. macOS용 Git 인스톨러는 Git 웹사이트에서 관리하고 있으며 <http://git-scm.com/download/mac> 에서 내려받는다.

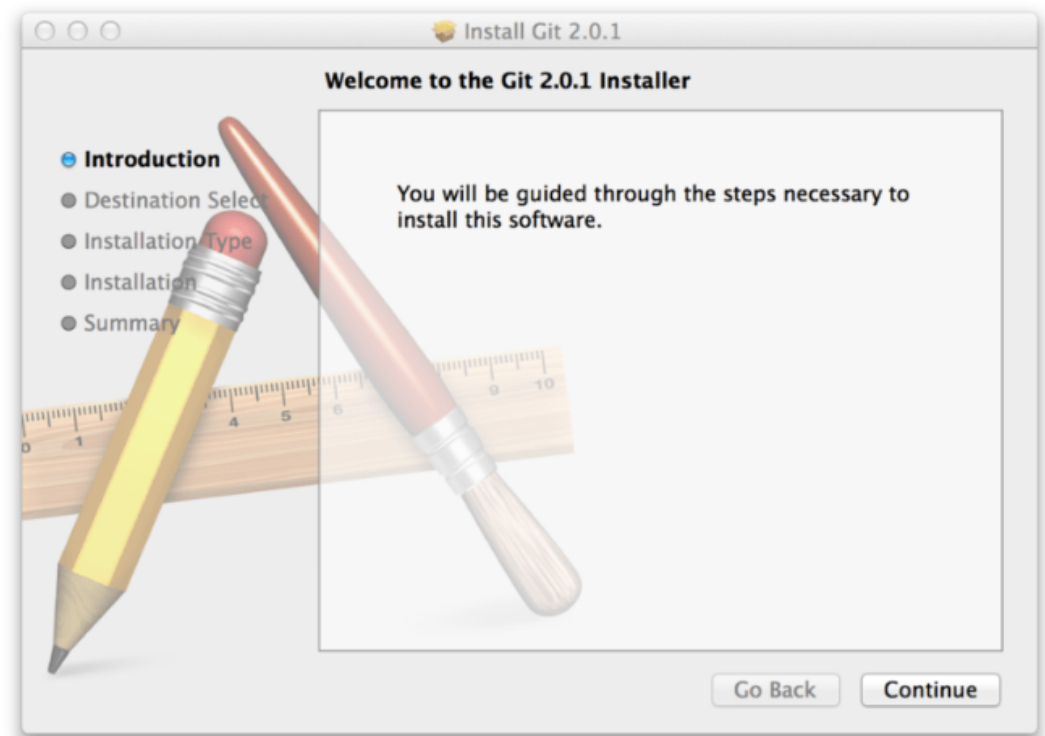


Figure 7. Git macOS 인스톨러.

'GitHub for Mac'을 설치하는 방법도 있다. 이 도구에도 CLI 도구를 설치하는 옵션이 있다. 'GitHub for Mac'은 <http://mac.github.com>에서 내려받는다.

2. Docker 설치 전 확인 사항

참고 <https://docs.docker.com/docker-for-mac/install/>

- Mac hardware는 2010 이후
- 인텔의 memory management unit (MMU) virtualization 지원
- 터미널에서 다음 명령으로 출력 확인

```
sysctl kern.hv_support  
kern.hv_support: 1
```

- macOS는 10.13 이상 (최신 버전으로 업그레이드 추천)
 - 4GB 이상 메모리
 - VirtualBox 4.3.30 이하 설치되어 있을 경우 삭제 요망
-
- 위 조건에 맞지 않을 경우 차선책으로 Docker Toolbox 설치 (아래 링크 참고)
<https://docs.docker.com/toolbox/overview/>
 - 또는 colab 환경에서 실습 진행 (매뉴얼 소개)

3. Docker 설치

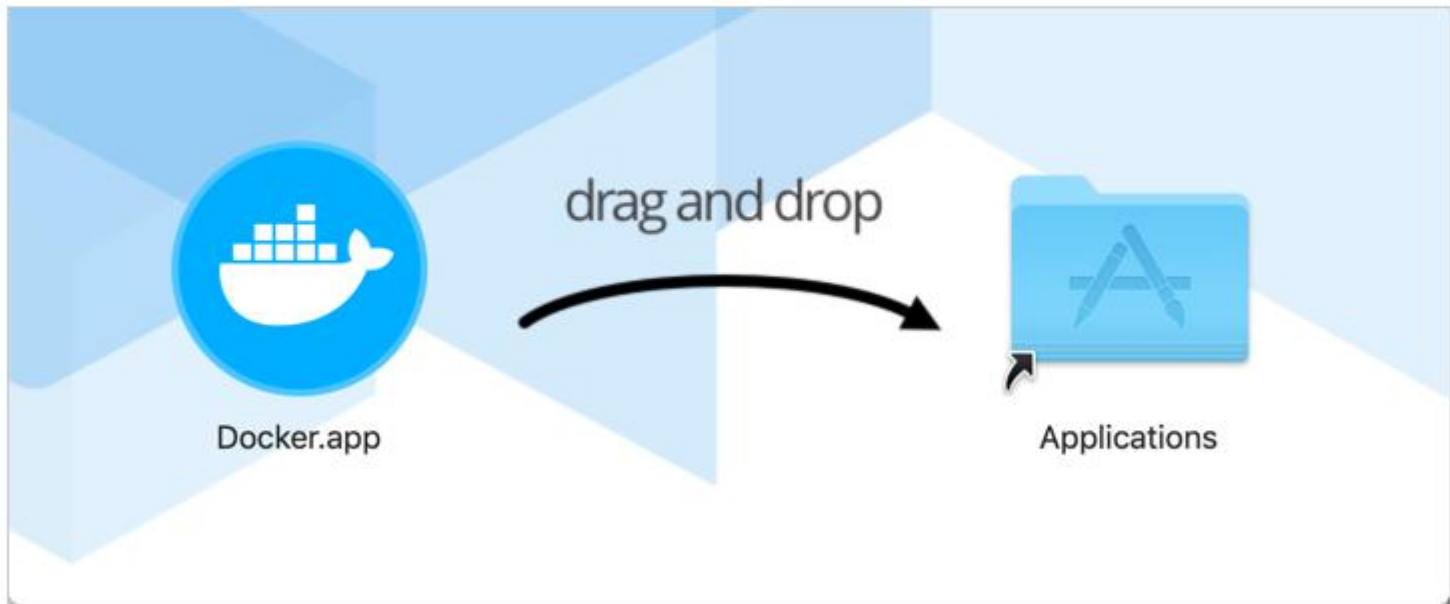
- Docker 설치 참고 <https://docs.docker.com/docker-for-mac/install/>
- (계정이 없을 경우) Docker Hub (<https://hub.docker.com/>) Get Started 클릭 후 사용자 등록
- (계정이 있을 경우) <https://www.docker.com> > Products > Desktop > Download Desktop for Mac and Windows 클릭 > 새 창 열리며 Sign in
- 빠른 시작 화면에서 설치파일 (Download Desktop for Mac) 다운로드
- 설치파일 실행 (설치 후 튜토리얼 계속 실습 추천)

3.1. Docker 설치

- <https://docs.docker.com/docker-for-mac/install/> 사이트 내용 참고

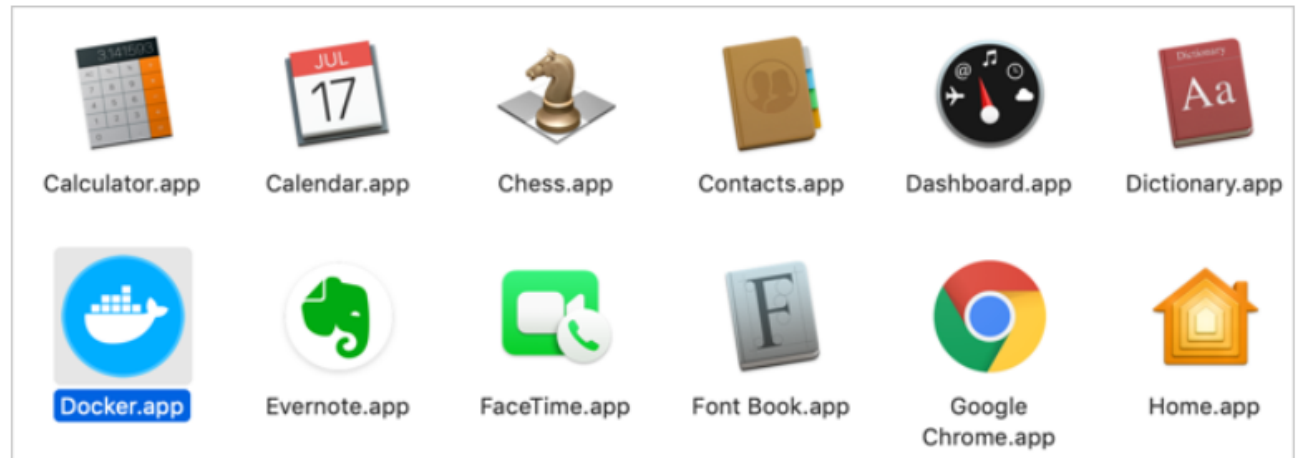
1. Install and run Docker Desktop on Mac

1. Double-click `Docker.dmg` to open the installer, then drag the Docker icon to the Applications folder.



2.

2. Double-click `Docker.app` in the Applications folder to start Docker. (In the example below, the Applications folder is in “grid” view mode.)

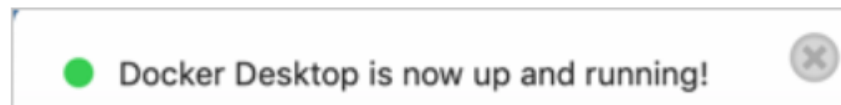



You are prompted to authorize `Docker.app` with your system password after you launch it. Privileged access is needed to install networking components and links to the Docker apps.

The Docker menu in the top status bar indicates that Docker Desktop is running, and accessible from a terminal.



If you just installed the app, you also get a message with suggested next steps and a link to the documentation. Click the Docker menu (🚢) in the status bar to dismiss this pop-up notification.



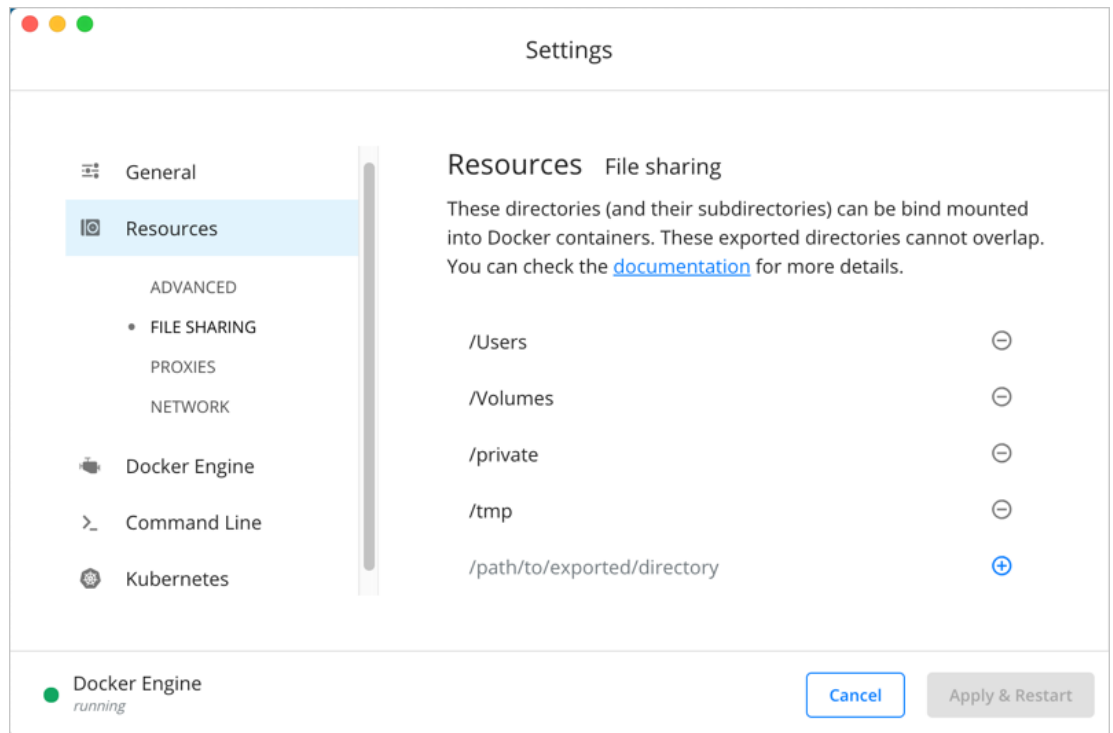
- 3.** 3. Click the Docker menu () to see **Preferences** and other options.
4. Select **About Docker** to verify that you have the latest version.

Congratulations! You are now successfully running Docker Desktop.

- 구글 검색을 통해 인터넷 문서를 참고하시기 바랍니다.

3.2 디렉토리 공유 설정

- 컨테이너 저장소는 휘발성으로 컨테이너 종료와 동시에 사라짐. 따라서 컨테이너 저장소와 하드드라이브 공유 필요
- <https://docs.docker.com/docker-for-mac/> 참고
- 도커 메뉴 > Preferences > Resources > File sharing
- 아래와 같이 타깃 디렉토리 (/tmp) 공유



4. 도커 이미지 다운로드

- 터미널 실행
- 다음 명령으로 필요 파일들 복사 및 이미지 다운로드 (약 5.3GB)

```
> cd /tmp/  
/tmp> git clone https://github.com/greendaygh/bioengml.git  
/tmp> cd bioengml  
/tmp/bioengml> docker pull haseong/bioengml:v1  
/tmp/bioengml> docker images
```

colab (코랩) 환경 설정

구글 colaboratory (colab)은 주피터 노트북을 구글 서버에서 무료로 사용할 수 있도록 제공하는 서비스입니다. 다양한 데이터분석 python 패키지들이 기본으로 설치되어 있으며 무료로 GPU를 사용할 수도 있습니다.

본 실습에서는 Docker 환경 설정이 되지 않는 분들을 위한 실습 환경으로 사용될 예정이며 다음 소개해 드리는 순서로 설정을 해 오시면 되겠습니다.

1. 지메일 계정 만들기

- <https://www.google.com/gmail/>
- 계정 만들면 google drive 사용 가능

내 파일을 내 마음대로 보관 - 드라이브 기능

15GB 저장 공간

어떤 파일이든 보관

공유도 내 맘대로

안전하게 보관

Google 계정이 있으면 15GB의 저장용량을 무료로 사용할 수 있습니다. 이 저장용량으로 Google 드라이브, Gmail, Google 포토에서 파일 및 이메일 첨부파일을 저장하고 사진과 동영상을 백업할 수 있습니다. 15GB 이상의 저장용량이 필요하다면 [Google One](#)으로 업그레이드하세요.

15GB
FREE



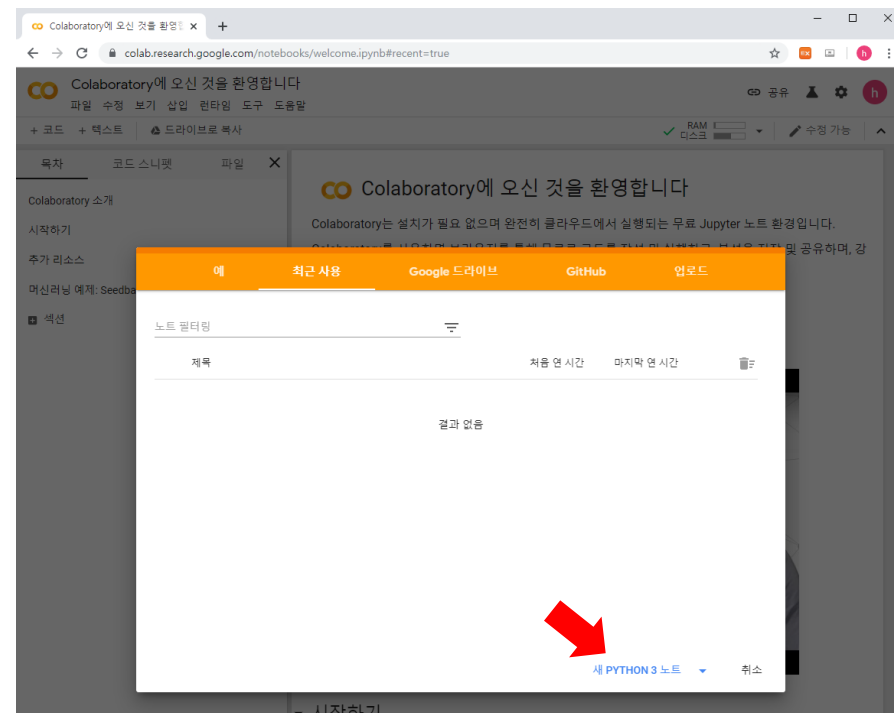
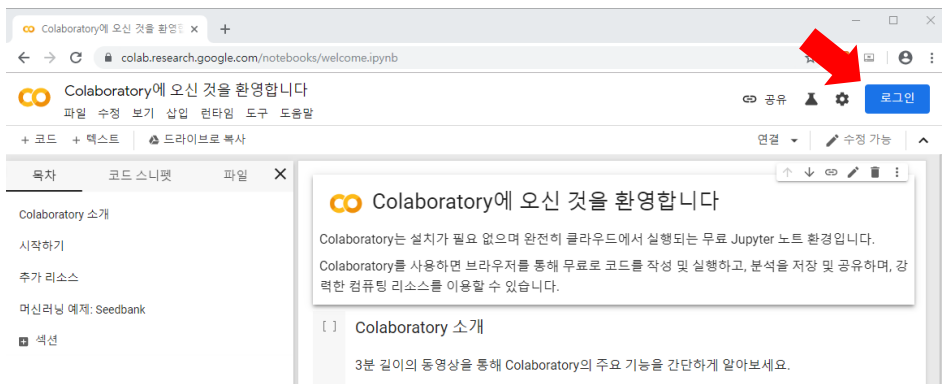
2. Chrome 브라우저 설치

- colab은 JupyterLab을 사용하는 환경으로 Internet explorer 지원이 안되며 Chrome 브라우저 설치가 필요합니다
- 설치는

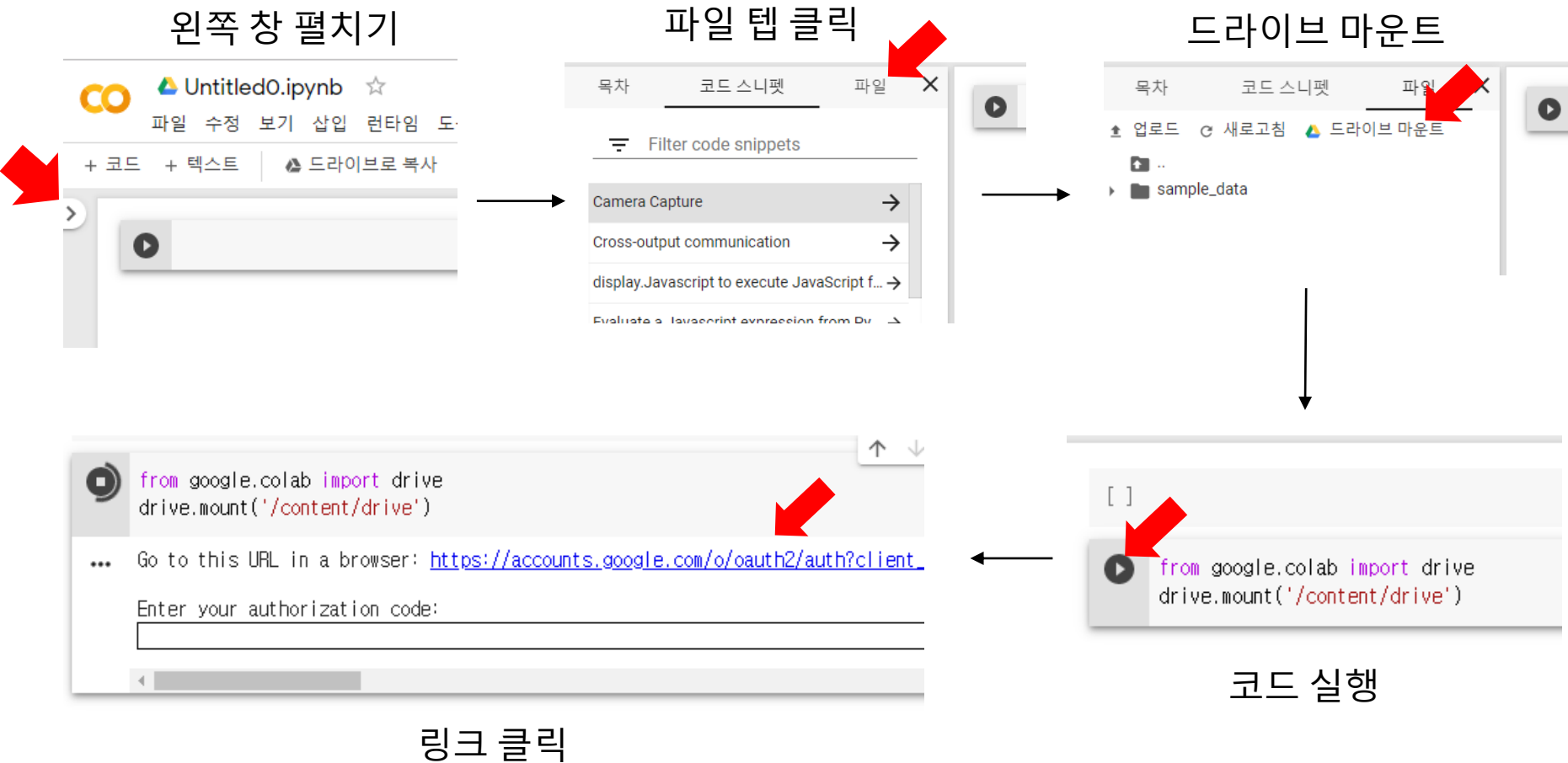
https://support.google.com/chrome/answer/95346?hl=ko&ref_topic=7439538# 이곳을 참고 하세요

3. colab 접속

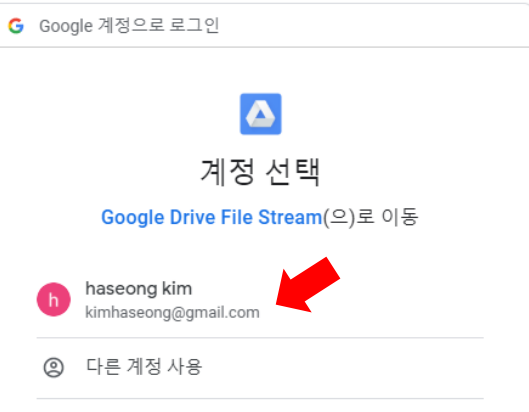
- <http://colab.research.google.com>
- 구글 계정으로 로그인
- 새 PYTHON3 노트 시작



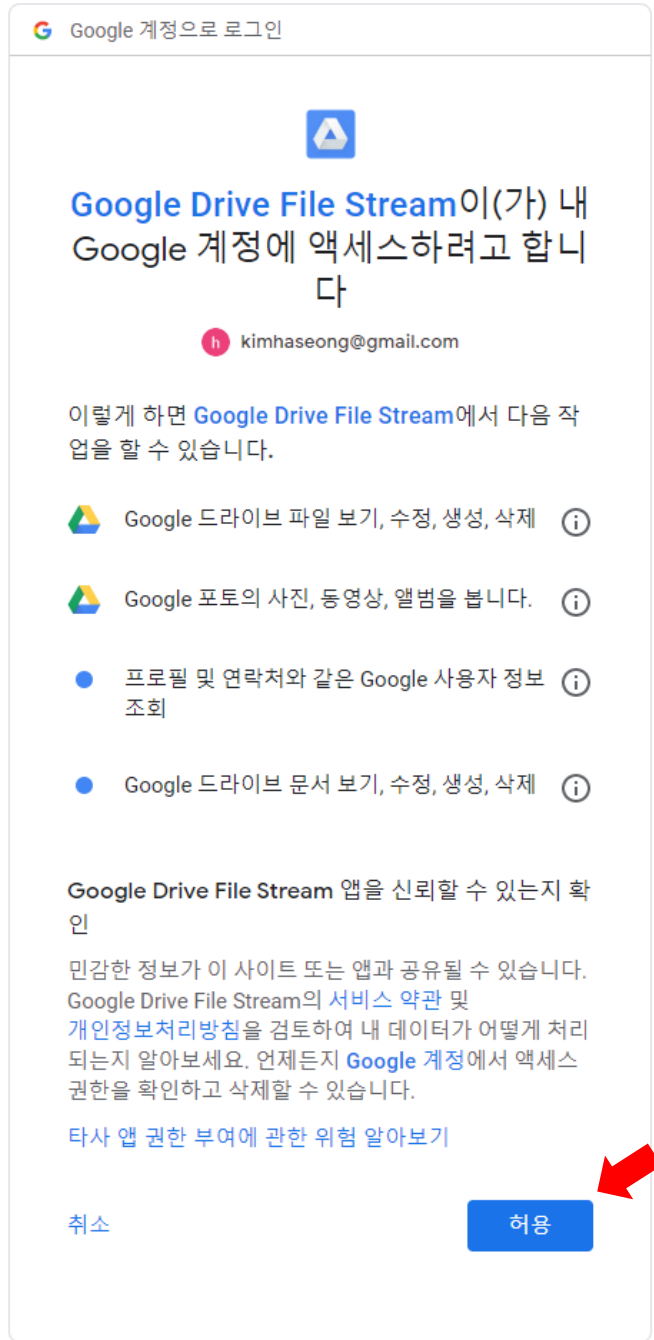
4. 구글 드라이브 연결



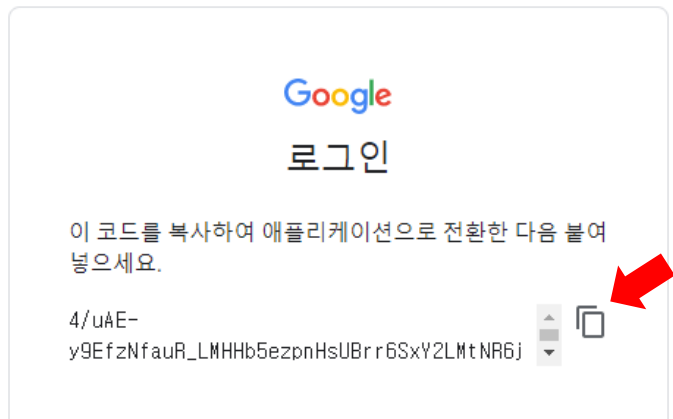
1.



2.



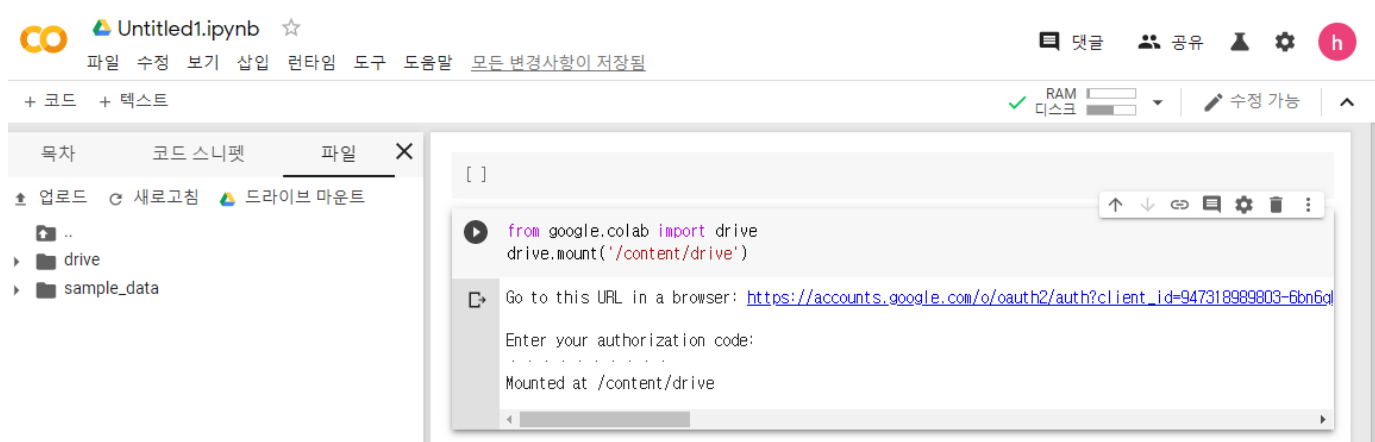
3.



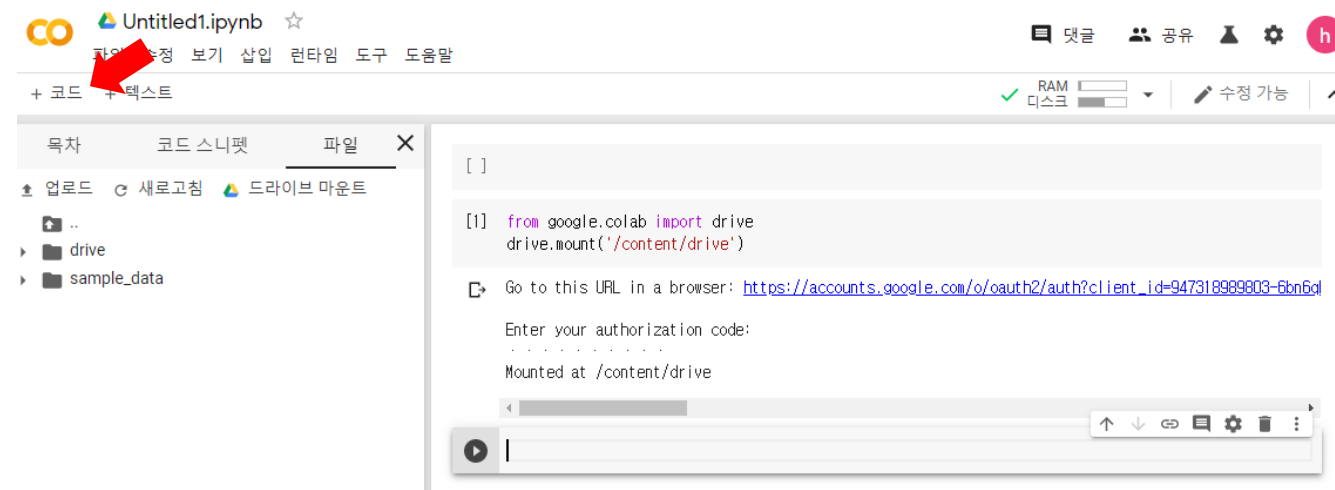
코드 붙여 넣고 Enter 입력



구글 드라이브와 연동 완료



+ 코드 클릭



5. 수업 내용 git 다운로드

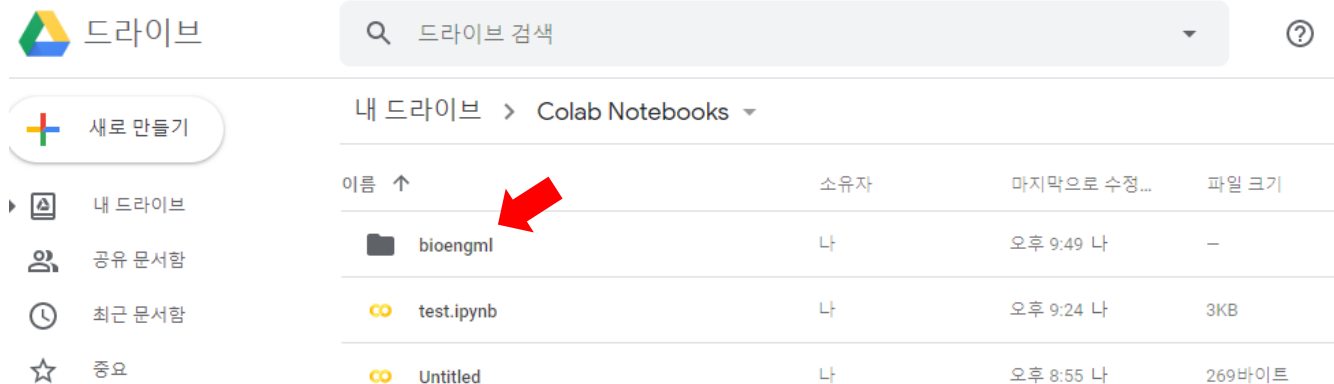
```
!git clone https://github.com/greendaygh/bioengml.git "/content/drive/My Drive/Colab Notebooks/bioengml"
```

코드 텍스트 박스에 위 명령 입력 후 Shift + Enter

```
[14] !git clone https://github.com/greendaygh/bioengml.git "/content/drive/My Drive/Colab Notebooks/bioengml"
```

```
Cloning into '/content/drive/My Drive/Colab Notebooks/bioengml'...
remote: Enumerating objects: 587, done.
remote: Counting objects: 100% (587/587), done.
remote: Compressing objects: 100% (378/378), done.
remote: Total 587 (delta 110), reused 572 (delta 97), pack-reused 0
Receiving objects: 100% (587/587), 4.68 MiB | 9.07 MiB/s, done.
Resolving deltas: 100% (110/110), done.
```

구글드라이브 Colab Notebooks 디렉토리에 bioengml 디렉토리 생성 확인



드라이브

새로 만들기

내 드라이브 > Colab Notebooks

| 이름 ↑ | 소유자 | 마지막으로 수정... | 파일 크기 |
|------------|-----|-------------|--------|
| bioengml | 나 | 오후 9:49 나 | - |
| test.ipynb | 나 | 오후 9:24 나 | 3KB |
| Untitled | 나 | 오후 8:55 나 | 269바이트 |

6. 구글드라이브에서 .ipynb 파일 열기

- colab 상단 메뉴 > 파일 > 드라이브에서 찾기
- 새창 열리면 bioengml 디렉토리 클릭
- bioengml-Day1.ipynb 파일 오른쪽 마우스버튼
- 연결 앱 > Google colaboratory 클릭

새로운 colab 창 열리면 “+코드” 클릭, biopython 모듈 설치

```
!pip install biopython
```

Docker Windows,
MacOS 공통

Jupyter lab 실행

