Google Colab for Machine Learning

By: Shuhao Lai

Date: 7/21/19

General Knowledge

There is a Google Drive folder automatically created for Google Colab. This folder stores all my notebooks. In order to directly access and manipulate files in that folder, I must "mount DRIVE".

The text field is rendered in mark down.

Use GPU by going to "runtime" → "Change runtime type" → select options.

Limits for GPU

Unlimited supply of 12 hours of continuous access to a k80 GPU.

Running Shell Commands

Note that the following rules apply to IPython, which is essentially Jupyter. Consequently, these rules apply to Google Colab, which is built on top of Jupyter.

Use "%%shell" to make the entire cell run as shell commands. Use "!command" to run that one line as a shell command. Both these methods run the shell commands on a new subprocess, not the current one in use. Thus, "!cd path" actually changes directory for a subprocess and has no effect on your process, which is likely what you want to change.

Use "%command" to affect your current process. Thus, "%cd path" has the expected effect. Further, "cd path" is shorthand for "%cd path"

Note: "!" might not run on shell. There is probably a more complicated method to choose bash or shell.

Uploading and importing .py File

Open a notebook in Google Colab, then click ">" → "Files" → "Upload" → upload your Python file. The file will be uploaded to root. This might not be ideal as you want the file to be in the same place as your notebook.

To add file to the same folder as the current notebook, simply upload the desired file wherever your notebook is stored in your drive. Now in your notebook, you can now import as you normally do; "import example_module".

Importing .ipynb File Already in Colab

To import existing ipynb files, you can export the file as a py file and follow the above instructions to import it as a module. However, you can also simply use the command "%run MyOtherNotebook.ipynb"

Mount Google Drive

Click ">" → "Files" → "MOUNT DRIVE". Code will be given to you to mount your drive. Your drive is only mounted on that one notebook.

Using with GitHub or Bitbucket

Default working directory is "/content". Useful to "%cd" to desired directory where you want clone to be stored. Then, execute "!git clone …" to clone repository to the current directory. The .py file will not be editable in Google Colab; however, they can be imported as modules if the notebook is in the same directory.

"Git clone" the provided link from your git host to clone a public repository. Use the following format to git clone a private repository:

!git clone https://username:password@github.com/username/repository.git

Extra

Shell is the command prompt provided by Windows and does not recognize "ls" command. Bash is "Bourne Again Shell" which recognizes "ls".

Default folder for a Notebook is "/content". This is usually where Drive is mounted to and where all files are uploaded to.