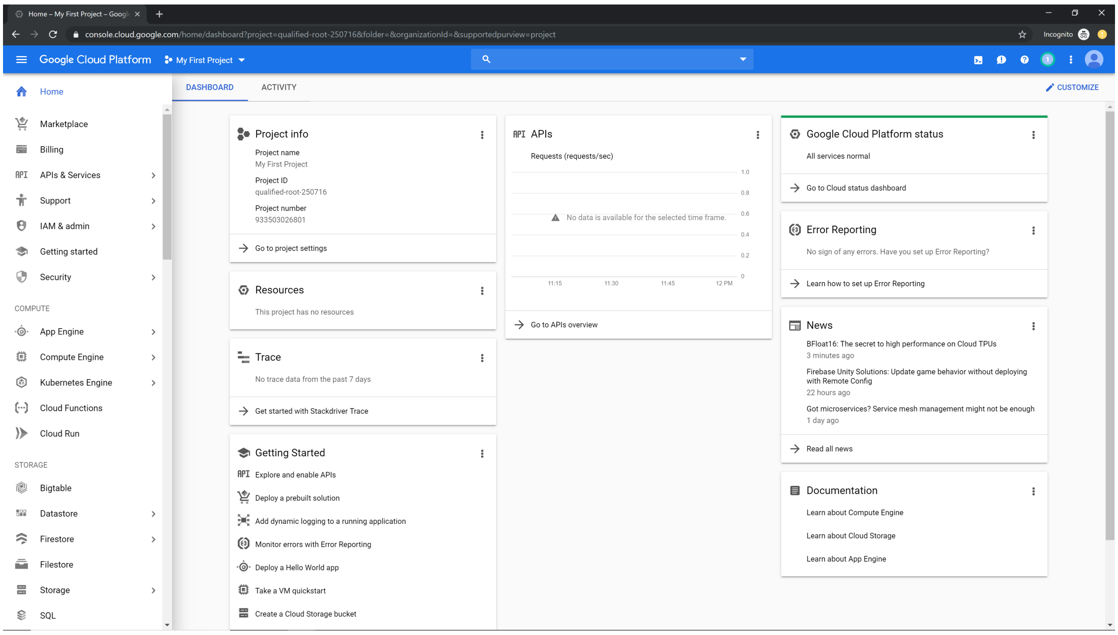
Google Cloud Platform Setup Instructions

Kathleen McKeown, Fall 2019
COMS W4705: Natural Language Processing

1 Google Cloud Platform Account

1.1 Account Setup

Begin by creating a Google Cloud Platform account here **using your LionMail account**. On a successful login you will see the following screen.



Project info

Project name
My First Project

Project ID
qualified-root-250716

Project number
933503026801

Go to project settings

Resources

This project has no resources

Trace

No trace data from the past 7 days

Get started with Stackdriver Trace

Getting Started

Explore and enable APIs

Deploy a prebuilt solution

Add dynamic logging to a running application

Monitor errors with Error Reporting

Deploy a Hello World app

Take a VM quickstart

Create a Cloud Storage bucket

APIs

Requests (requests/sec)

No data is available for the selected time frame.

Go to APIs overview

Google Cloud Platform status

All services normal

Go to Cloud status dashboard

Error Reporting

No sign of any errors. Have you set up Error Reporting?

Learn how to set up Error Reporting

News

BFloat16: The secret to high performance on Cloud TPUs
3 minutes ago

Firebase Unity Solutions: Update game behavior without deploying with Remote Config
22 hours ago

Got microservices? Service mesh management might not be enough
1 day ago

Read all news

Documentation

Learn about Compute Engine

Learn about Cloud Storage

Learn about App Engine

Google Cloud resources are contained in Projects; you may use the default “My First Project” or create a new Project by clicking on the project dropdown menu (at the top left, to the right of the main navigation bar) and selecting “New Project”.

1.2 Optional: Redeeming Coupons

When you create a new account, you are eligible for $300 of free credit, an offer which will be displayed on a banner at the top of your GCP homepage. However, in the event that you need additional funds and obtain a coupon code from the teaching staff, you may redeem it here at the “Education grants” page. This can be done at any time, even if you have redeemed other coupon codes in the past.

Make sure that you are on the correct account when redeeming coupon codes! You may wish to do this in an incognito window to ignore any other Google logins you may have.

2 Virtual Machine Setup

2.1 SSH Keys

Google Cloud virtual machines use a type of access credential called SSH keys to allow you to connect. These keys consist of a public-private encryption key pair, where the two keys are linked mathematically. The **public key**, which you provide to Google Cloud, allows any user to encrypt a message that can then be decrypted with the **private key**, which you keep to yourself. This allows Google Cloud to verify your identity when connecting. Google Cloud takes care of this when you connect using its channels (e.g., through the browser SSH), but since you will sometimes want to do GUI tasks that those channels cannot handle (e.g., plotting a graph or visualizing weights), you will need to set up your own key pair to use elsewhere.

2.1.1 Terminal Setup

First, set up the terminal you will use to connect so that you can perform the aforementioned GUI operations (called X11 forwarding).

Windows Users: We recommend you download MobaXTerm, which is capable of X11 forwarding. When you wish to connect to your VM, start MobaXTerm, start a local terminal, and connect as usual. Note that MobaXTerm uses Linux-style commands.

Mac Users: Edit the `sshd_config` file (`sudo vim /private/etc/ssh/sshd_config`, or `sudo vim /private/etc/sshd_config`), and change X11Forwarding to yes (uncomment if commented). If you do not have XQuartz installed, install that for X11 forwarding. Also make sure that you use `ssh -Y uni@instance_ip`, instead of just `ssh uni@instance_ip`, when trying to logon to the instance, as mentioned below.

2.1.2 Creating the SSH Keys

1. Open a terminal window and type the following two commands:

```
ssh-keygen -t rsa -f ~/.ssh/my_w4705_key -C [MY_UNI]
chmod 400 ~/.ssh/my_w4705_key
```

where of course MY_UNI is your UNI, or any username you wish to use on your virtual machine.

The `ssh-keygen` command will create a private key file at `~/.ssh/my_w4705_key` and a corresponding public key file at `~/.ssh/my_w4705_key.pub`. The `chmod` command restricts the permissions on the private key—which must be kept secret — on your local machine so that only your user (and/or root) can view it.

1.5 (**Windows**). Windows users on MobaXTerm will need to start the ssh agent before running the next step. Execute

```
eval ‘ssh-agent -s’
```

If after running this command, you receive the error “Could not open a connection to your authentication agent.” on running the next step, execute the following:

```
ssh-agent -s
```

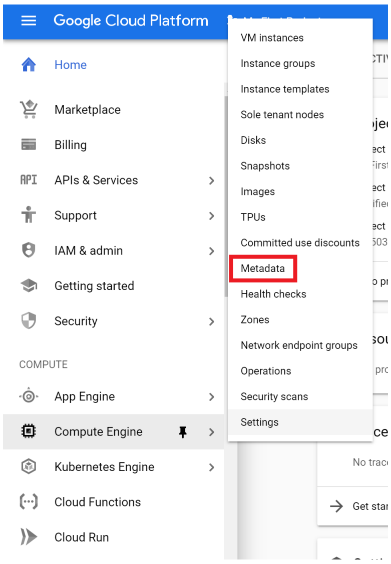
And then copy-paste the output of the above command into your command prompt. You will need to start the `ssh-agent` every time you wish to connect to your virtual machine.

2. (**All**) You must also add the private key to your `ssh-agent`:

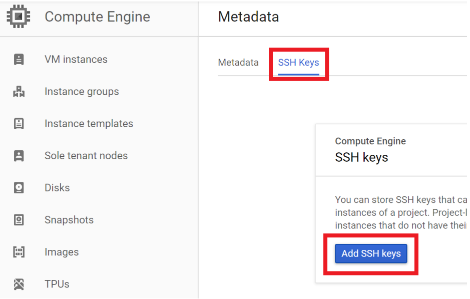
```
ssh-add ~/.ssh/my_w4705_key
```

2.1.3 Adding SSH Keys to Your Project

To add the public SSH key to your Gogle Cloud project, first navigate to “Compute Engine”, then “Metadata” using the navigation bar on the left side of the screen. You may need to wait here while Compute Engine starts up for the first time.



Select “SSH Keys” and then ”Add SSH Keys”.



You can store SSH keys that ca
instances of a project. Project-l
instances that do not have thei

Paste the entire contents of your public key file (`~/.ssh/my_w4705_key.pub`) into the text box (you can open the `.pub` file in a text editor to get its contents). The key you paste should begin with "ssh-rsa". When you paste your public key, it automatically fills in the "Username" field, which is the login username associated with the public key. You can change the "Username" field by modifying the very last part of the public key that you pasted. Make sure to save the SSH key when you are satisfied.

2.2 VM Creation

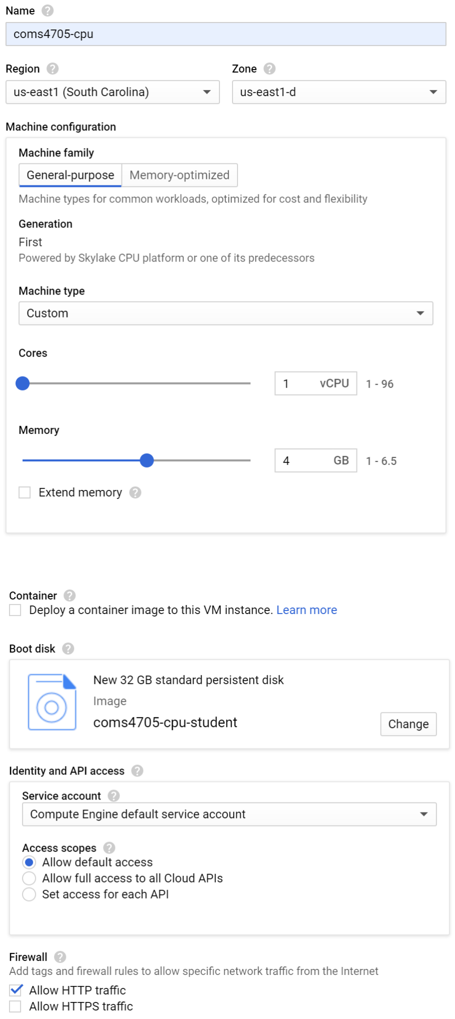
Finally, create a virtual machine to use for your programming assignments. The TAs have given you access to a premade VM image on Google Cloud Platform, assuming you are using your Lionmail account. If for some reason you do not have access, **you must contact the TAs immediately and provide the email you used to sign up for Google Cloud**. The TAs will use this image to grade Homework 0 and any assignments which do not involve GPUs, so it is to your benefit to use it for writing or testing your code.

Navigate to the "VM instances" page under "Compute Engine" (the same place you found the metadata menu in Step 3.1.3) and click "Create".

2.2.1 Virtual Machine Specifications

We provide the specifications we used for our provided VM image; we recommend that you set up your machine the same way, although it is not required.

You can use as much processor memory and as many cores as you like; pay attention to the estimated cost of the instance shown to the right of the menu. Make sure to check "Allow HTTP traffic" so that you can log in to your instance. You do not need to change any advanced options or add any startup scripts or GPUs, although you can. You may wish to inspect the availability options under "Management, security, disks, networking, sole tenancy"; for example, turning off "Automatic restart" may help save money and show that there was a problem if your VM shuts down unexpectedly.



Name
coms4705-cpu

Region: us-east1 (South Carolina)
Zone: us-east1-d

Machine configuration

Machine family: General-purpose | Memory-optimized
Machine types for common workloads, optimized for cost and flexibility

Generation
First
Powered by Skylake CPU platform or one of its predecessors

Machine type: Custom

Cores: 1 vCPU (1 - 96)

Memory: 4 GB (1 - 6.5)

□ Extend memory

Container
□ Deploy a container image to this VM instance. Learn more

Boot disk
New 32 GB standard persistent disk
Image
coms4705-cpu-student

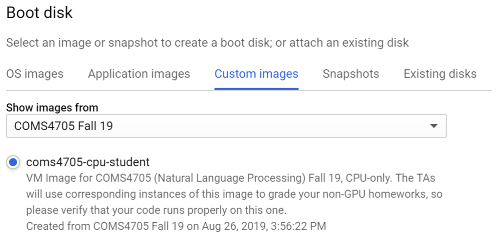
Identity and API access

Service account: Compute Engine default service account

Access scopes
● Allow default access
○ Allow full access to all Cloud APIs
○ Set access for each API

Firewall
Add tags and firewall rules to allow specific network traffic from the Internet
☑ Allow HTTP traffic
□ Allow HTTPS traffic

In order to select the premade VM image, select "Change" next to the boot disk, select "Custom Images", and locate "coms4705-cpu-student" as below; make sure to check it and click "Select".



Boot disk

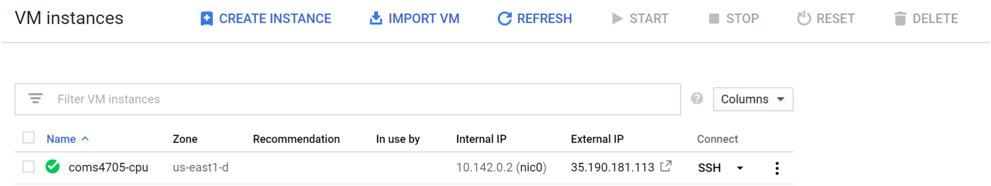
Select an image or snapshot to create a boot disk; or attach an existing disk

Show images from: COMS4705 Fall 19

● coms4705-cpu-student
VM Image for COMS4705 (Natural Language Processing) Fall 19, CPU-only. The TAs will use corresponding instances of this image to grade your non-GPU homeworks, so please verify that your code runs properly on this one.
Created from COMS4705 Fall 19 on Aug 26, 2019, 3:56:22 PM

Finally, click "Create" to create your instance. The VM will take some time to start up and you will be redirected to the previous "VM instances" page.

2.2.2 Logging In



□	Name	Zone	Recommendation	In use by	Internal IP	External IP	Connect
□	✓ coms4705-cpu	us-east1-d			10.142.0.2 (nic0)	35.190.181.113	SSH

You can access your VM via SSH, similar to how you access a CLIC machine via SSH. Some important points:

1. The username you use must be the "Username" that you entered for your SSH key in Section 3.1.3.

2. The IP address you connect to is the "External IP" for your VM instance. This IP will change every time you restart your instance, so make sure to check it each time.

3. You must use the private key that you created in Section 3.1.2.

Make sure your VM is on (the green check appears on the "VM instances" page as above) and connect using SSH:

Windows: `ssh <your_UNI>@<your_external_IP>`
Mac: `ssh -Y <your_UNI>@<your_external_IP>`

Note that <your_UNI> is the username that GCP parsed out from the public key that was pasted in Section 3.1.3, and that <your_external_IP> is the external IP address shown on your VM instances page (so in the example image above, 35.190.181.113).

For generic SSH client details, see GCP's guide here. For quick access, you can also use GCP's browser SSH client which you launch simply by clicking on the SSH button on the "VM instances" console. However, do note that this SSH client cannot perform X11 forwarding, which you may want or need to view plots, etc.

2.2.3 Shutting Down

Remember to shut down your VM instance promptly when you are done using it, as you are charged as long as the VM instance is running and may use up your coupon credit! Shut the VM down by checking its box on the "VM instances" page and clicking "Stop" on the top bar, or selecting "Stop" from the drop-down menu beside the VM (the three dots in the image in Section 3.2.2). The VM can be started equivalently (selecting "Start" in either location).

2.3 Python on the Virtual Machine

2.3.1 conda Setup

Your virtual machine is now ready! The only thing you will likely wish to do is edit your personal .bashrc file to tell it where Anaconda, which is a manager for Python, lives. To do that, simply log in to your VM and run the following from your home directory:

```
/opt/anaconda/bin/conda init
source ~/.bashrc
```

Now you will be able to use conda instead of /opt/anaconda/bin/conda from the command line.

2.3.2 Using Anaconda

Anaconda is a package and version manager for Python, the language in which you should expect to code for this class. We have installed Anaconda on our premade VM images and set up a basic environment with basic required packages. While Anaconda and this environment are not required, we will test your homework using the same environment, so it is to your benefit to use them for development or testing.

This environment includes **Python 3.6.9**. **We WILL NOT accept code that does not run in Python 3.6.9, INCLUDING because of** Python 2 print errors.

Our VM images have Anaconda installed at /opt/anaconda with a premade coms4705 environment at /opt/anaconda/envs/coms4705. You can activate this environment using conda activate coms4705 from any location on your VM and deactivate it with conda deactivate. Python packages can be installed through conda as well if you wish to use a package we have not already installed (but remember to include it in a requirements.txt file with your homework).

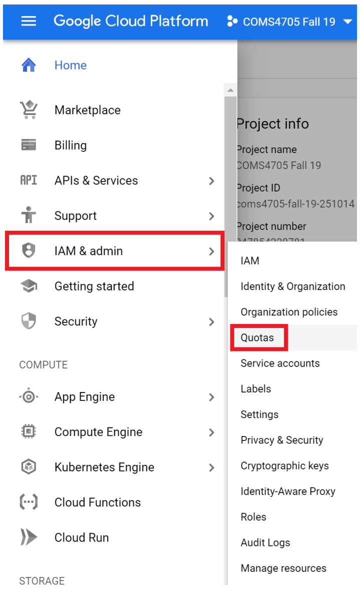
The packages we include by default are gensim, lxml, matplotlib, nltk, numpy, pandas, scikit-learn, scipy, and seaborn and unless otherwise specified you may always use these libraries for any task.

You may wish to read up on the conda documentation here.

2.4 GPU Machines

For later homework assignments, you will want to use a GPU-enabled machine; however, GPUs cost quite a lot to operate, so we initially provided the previous CPU-only image, and you may wish to continue using it when you are debugging or writing non-GPU code. In order to use GPUs on a virtual machine, you must first request a quota increase (i.e., you must request permission from Google Cloud to use more GPUs).

Navigate to “IAM”, then “Quotas” from the main navigation bar.



Project info

Project name
COMS4705 Fall 19

Project ID
coms4705-fall-19-251014

Project number

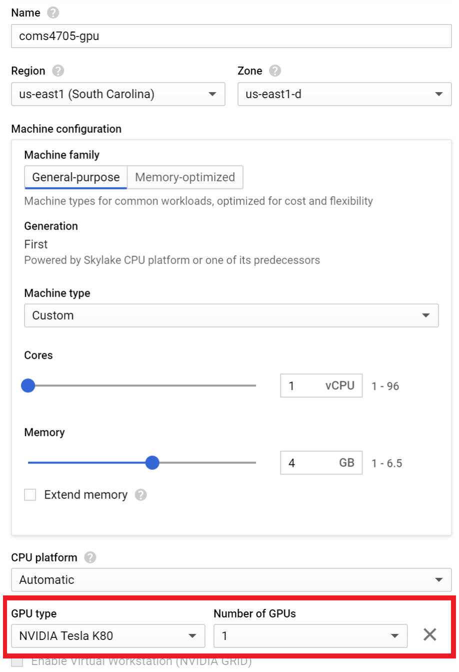
Select “Edit Quotas” and check “GPUs” from the quota list, then request the increase (you will not need more than one GPU). You should receive an automated email a short time later confirming your quota increase, at which time you may create and use VM instances with GPUs.

2.4.1 Quotas

2.4.2 Virtual Machines with GPUs

When GPU-requiring homeworks are released, the TAs will make available to you, in the same way as before, another VM image with relevant software installed (`cuda`, `pytorch`, etc.).

Create a virtual machine the same way as before, with these added GPU specifications.



Name

coms4705-gpu

Region: us-east1 (South Carolina)

Zone: us-east1-d

Machine configuration

Machine family: General-purpose

Machine types for common workloads, optimized for cost and flexibility

Generation: First

Powered by Skylake CPU platform or one of its predecessors

Machine type: Custom

Cores: 1 vCPU 1 - 96

Memory: 4 GB 1 - 6.5

□ Extend memory

CPU platform: Automatic

GPU type: NVIDIA Tesla K80

Number of GPUs: 1

□ Enable Virtual Workstation (NVIDIA GRID)

In order to use Tesla K80 GPUs, you MUST set your region to us-east-1 (South Carolina) and your zone to us-east-1d! These are the least expensive type of GPU offered by Google Cloud and we recommend you use one. None of our assignments will require anything more powerful. We will not provide additional coupons for more and more powerful GPUs.

The GPU image is exactly the same as the first, CPU-only image in all other respects (`conda` environment, etc.). The TAs will use this image to grade assignments involving GPUs. Any code that runs on the first image will also run on this image (assuming you list your additional requirements in a `requirements.txt` file).