**Introduction**

The cloud classification project is implemented in two methods:

1. Standard CPU based - "cloud\_combine.py" script is to train the DNN model for detecting cloudy pixels of Cross- track Infrared sounder using the top 75 Principal Components (PC)of it's spectrum.

2. GPU based

**Software requirements**

1. Python 3.7 or above

2. Python IDE (Visual Studio Code)

**Standard CPU-based implementation**

**I - Clone the repository -** [**Link**](https://github.com/stccenter/ArcCI_DL.git)

**II - Set up the virtual environment**

1. Create a new folder and name it as cloudclassify.

2. Copy cloud\_subpixel.py from cloned repository and place it inside cloudclassify folder.

3. In Visual Studio Code, go to Terminal and run the below in cmd terminal. This creates a virtual environment called "cloudclassify-venv"

python -m venv cloudclassify-venv

4. For Windows, run below line to activate the virtual environment

cloudclassify-venv\Scripts\activate.bat

**III - Install python packages**

Use the following command to install all necessary packages at once:

pip install -r requirements.txt

**IV Download the data**

1. Click the [link](https://drive.google.com/drive/u/0/folders/1d9uS1EDtIkmTHu3pDhJgR7mbqVS2gZNM) and download the input data.

2. Place the downloaded "cloud\_pc.sav" file inside cloudclassify folder

**V - Run the script**

Now, we are all set to run the script. Run the cloud\_combine.py script using below command. This script accepts anargument called flag (-f). The default value of the flag is detection.

1. Run the script with default value "detection"

python cloud\_combine.py

2. Run the script with flag "rainy cloud"

python cloud\_combine.py -f "rainy cloud"

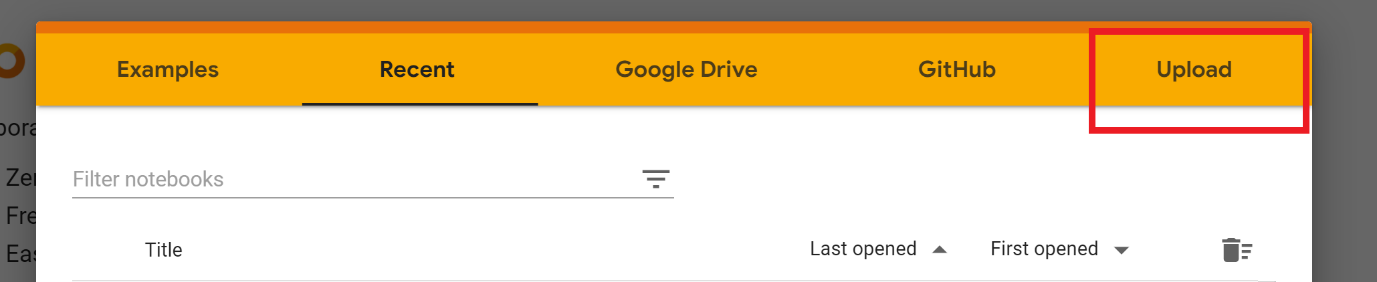
**VI - Output**

1. Once the script is finished running, you will find the output inside my\_model folder.

**GPU-based implementation**

**I - Open Google Colab**

1. Go to [Google Colab](https://colab.research.google.com/notebooks/intro.ipynb#recent=true).

2. Click on the Upload

3. Choose the "CloudClassifyGPU.ipynb" downloaded from the Git repository.

**II - Download train and test data and upload them in your Google Drive**

1. Download data using this [link](https://drive.google.com/drive/folders/1XqrxJd6rGgd0N2QJXRd1fm5aCZiSZClR?usp=sharing).
2. Download all three folders to your local machine.
3. Then upload it to your Google drive.

4. Create a new folder called "logs" inside the same folder where you uploaded the above three folders.

5. Finally, your folder should have four folders. See screenshot below:

Text

Description automatically generated

**III - Change Runtime to GPU**

1. Go to Runtime in menu.

2. Click "Change runtime type". image

3. In Notebook setting, change Hardware accelerator to "GPU" and click Save.

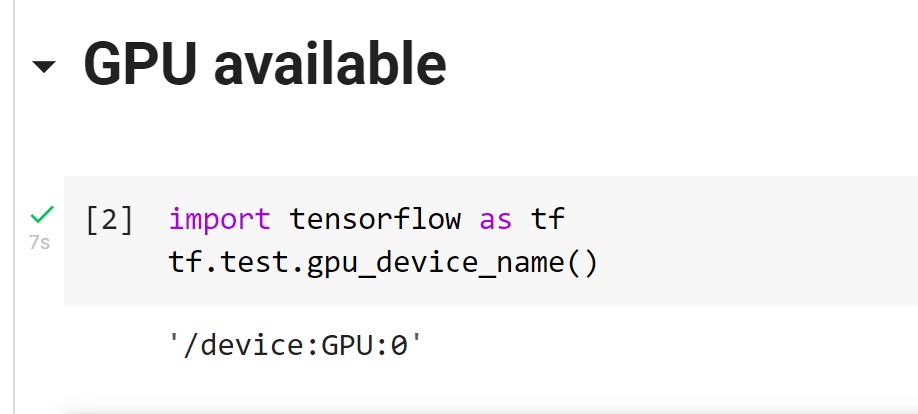
Graphical user interface, text, application

Description automatically generated

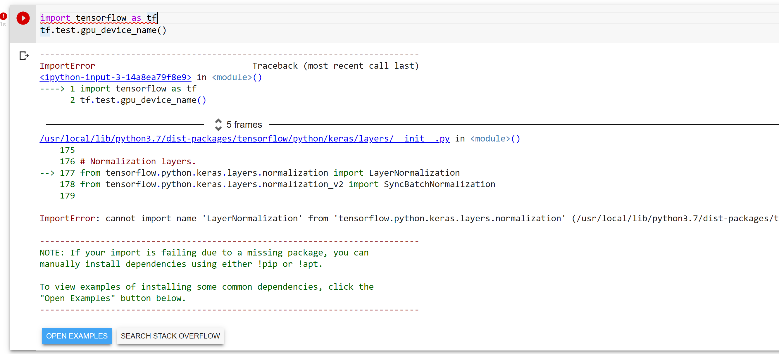
**IV - Run the notebook**

1. Run each code block sequentially.

2. Note when you run the "GPU available" block, you should see below if GPU is properly enabled.

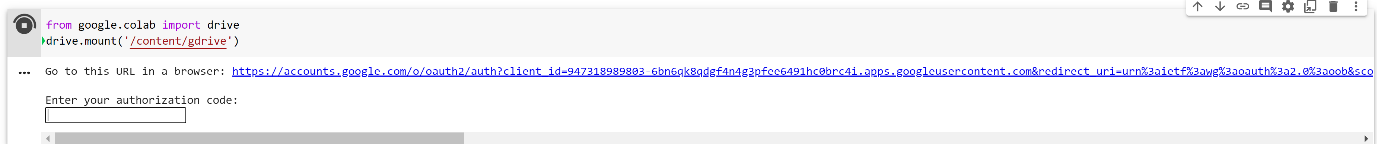


Note: If you encounter Import error (screenshot below): cannot import name LayerNormalization, go to Runtime -> Factory Reset runtime. Then set "Hardware Accelerator" to GPU and save it.



3. When you run the code block "Mount the Google drive".

1. Click on the URL.



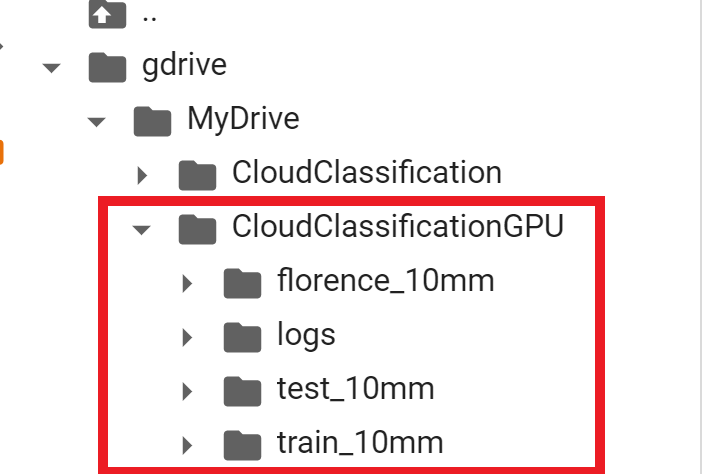
2. Obtain the authorization code and paste it in the "Enter your authorization code" box.

3. Verify Google drive is mounted correctly.

1. Click file icon in the left panel.

2. Click Mount Drive (third option).

3. Once it is mounted, you should see drive folder. See screenshot below:



1. Graphical user interface, text, application

   Description automatically generatedPlease make sure to check your data path if you get an IOError in the below code block.

**V - Output**

1. You will find the output inside the folder "my\_model"