

Lab Assignment #2 – Machine Learning on the cloud

Due Date: Friday, Week 12

Purpose:

The purpose of this Lab assignment is to:

1. To build (train), deploy and test a machine learning model, on the cloud.

General Instructions:

Be sure to read the following general instructions carefully:

1. This assignment must be completed individually by all the students.
2. Only provide the requested screenshots and make sure to have a complete screenshot, partial screenshots will not earn any marks.
3. You will have to add all the analysis and screenshots in the analysis report.
4. You will have to provide a **demonstration video for your solution** and upload the video together with the solution on **eCentennial** through the assignment link. See the **video recording instructions** at the end of this document.
5. In your 8-minute demonstration video you should explain your solution clearly, going over the main code blocks and the purpose of each method also demoing the execution of the code. Youtube links and links to google drive or any other media are not acceptable, the actual recording file must be submitted. Any submission without an accompanying video will lose 25% of the grade.
6. Any submission without an accompanying Analysis report will lose 100% of the grade.

Assignment Pre-requisites:

1. AWS school account credential
2. Amazon Sage Maker role
3. S3 bucket
4. Amazon Sage Maker notebooks

Exercise #1: Train a machine learning model using Amazon Sage Maker (100%)

In this exercise you will train a model to detect breast cancer using amazon Sage Maker. This is similar to the lab exercise in week #8.

You will carry out the steps listed in the example published by aws, with a few modifications explained in the exercise requirements section , the details are at this link:

https://sagemaker-examples.readthedocs.io/en/latest/introduction_to_applying_machine_learning/breast_cancer_prediction/Breast%20Cancer%20Prediction.html

Important :

Before you start, please do the following:

- 1) Read the steps in detail.
- 2) Read the required changes for this exercise, listed under exercise requirements.
 - a. Log into your aws account
 - b. Goto amazon SageMaker
 - c. Make sure the region is set to **Central**, as per below:



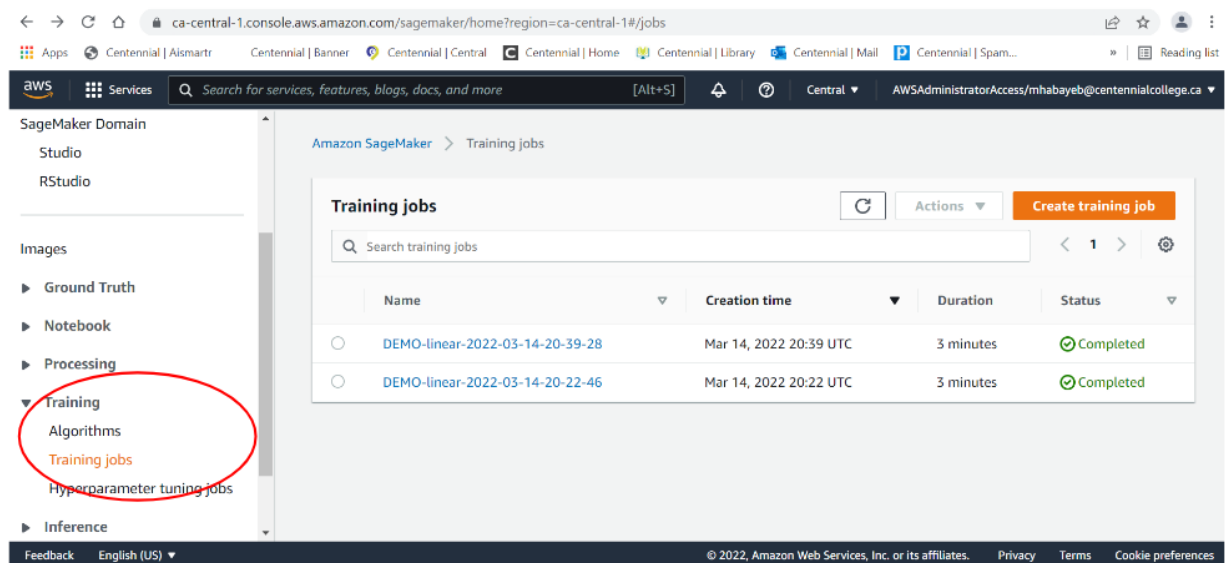
- 3) After completing the exercise make sure that you carry out all the necessary steps to **cleanup**, these involves:
 - a. Deleting the endpoint.
 - b. Deleting the items in the S3 bucket (data and model)
 - c. Stopping the notebook instance.
 - d. Deleting the notebook instance.

If you don't then you will incur further charges, to your account.

Exercise requirements

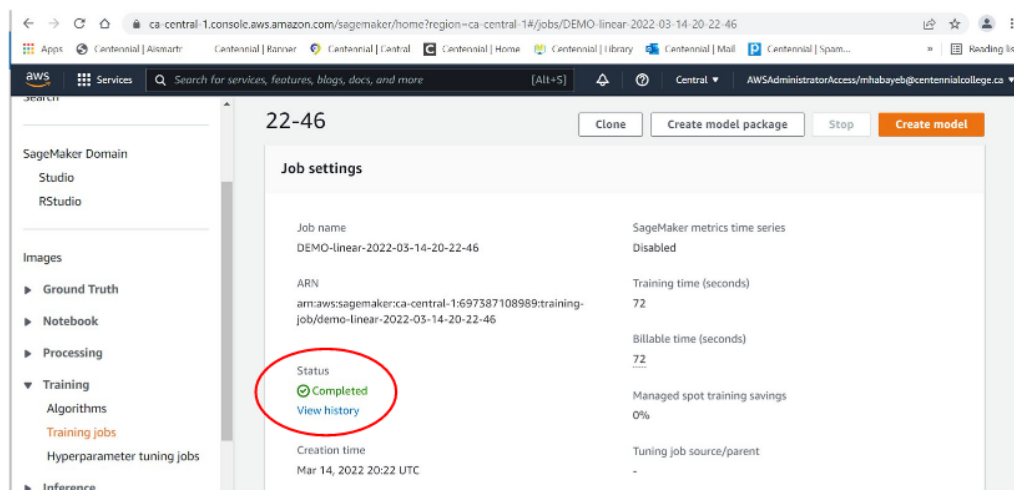
1. Log in to your aws account using the AWS SSO user portal i.e. not root.
2. Navigate to aws SageMaker
3. Create a new notebook instance and name it **firstname_breastcancer**
4. Once the instance is in service take a screenshot, name it **screen1_firstname** and add it to your analysis report.
5. Click on Jupyter and create a new notebook name it **firstname_ex1**
6. Create two cells, copy and paste the code under the section "setup" in the lab into your new notebook.
7. Change the value of the prefix from "sagemaker/DEMO-breast-cancer-prediction" to "sagemaker/DEMO-**firstname**-breast-cancer-prediction" where first name is your first name.
8. Execute both cells.
9. **Add** a cell to printout the bucket name, region name and the role name. Execute the cell, take a screenshot showing the output, name it **screen2_firstname** and add it to your analysis report.

10. Continue by copying and pasting & executing all the cells till you reach the “train” section. In the cell where you define the `linear_training_params` change the value of “`ephocs`” from 10 to 3 and change the value of “`num_models`” from 32 to 3. Then execute the cell.
11. Continue till you reach the “Cleanup” section.
12. **Important** - Delete the endpoint.
13. Download your notebook as html and attach it to the submission.
14. Download your notebook as a notebook .ipynb and keep a copy for you.
15. **Important** - Stop the notebook instance and then delete the notebook instance once stopped.
16. In SageMaker, navigate to “Training” and select “Training jobs” as illustrated in the below screen:



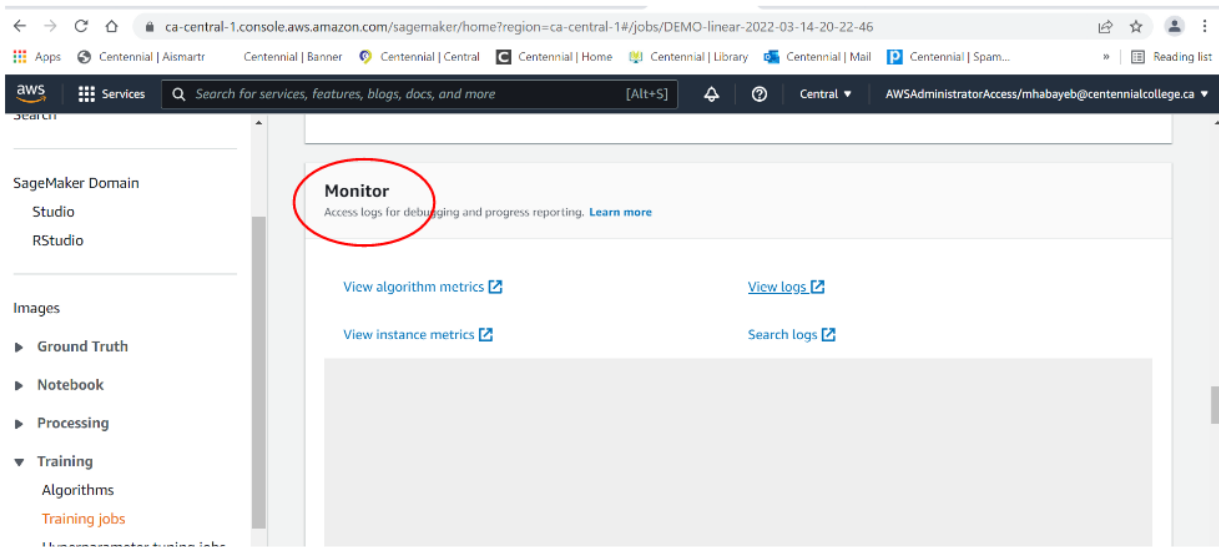
In your case you will have one training job, take a screenshot name it `screen3_firstname` and add it to your report.

17. Click on your training job, then click on view history, as illustrated in the below screen:



take a screenshot name it [screen4_firstname](#) and add it to your report.

18. Scroll down till you reach the “Monitor” section as per the below screen:

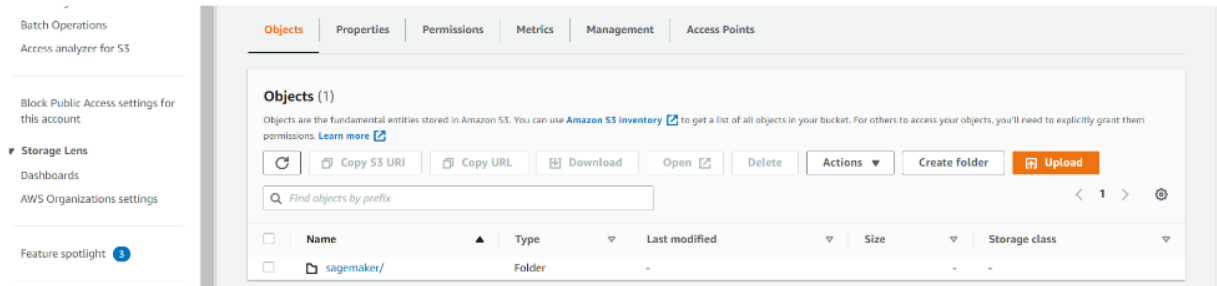


Click on view logs:

- Note in your analysis report which amazon service is launched.
- Take a screenshot of the [log event](#). Name it [screen5_firstname](#) and add it to your report.

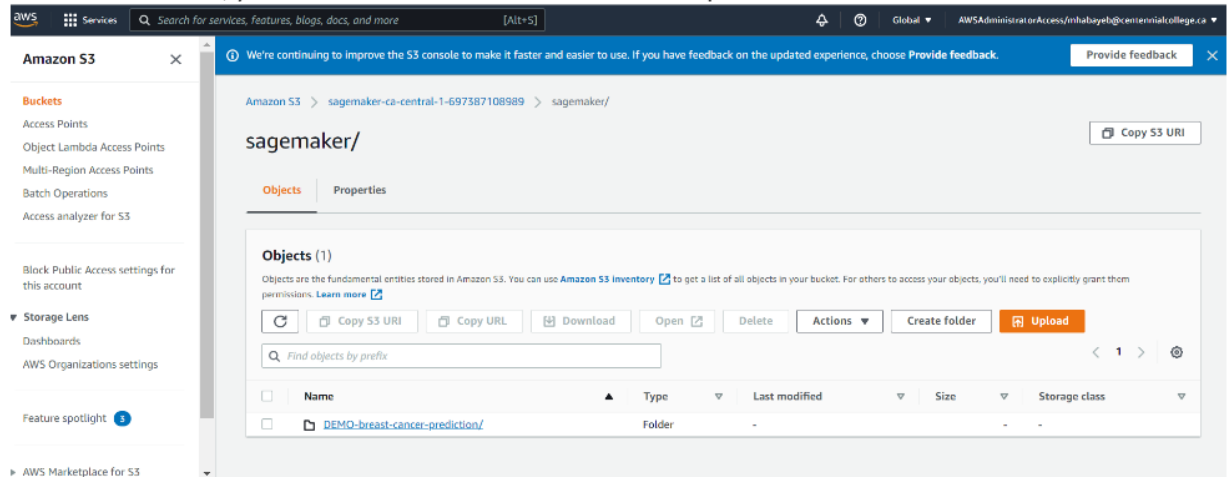
19. Navigate to your S3 services, check the newly created bucket that sagemaker created, click on it you should have a screen as illustrated below:

20.



take a screenshot of the log event. Name it [screen6_firstname](#) and add it to your report.

21. Click on the folder, you should find a folder for the demo as per the below screenshot:



22. Click on the demo folder, then take a screenshot. Name it **screen7_firstname** and add it to your report.

23. **Important** - Once you're done delete the Demo folder i.e empty the sage maker bucket.

24. In this exercise you used the SageMaker built-in "linear learner algorithm" research this algorithm. in your analysis report explain how this algorithm works, be thorough.

Naming and Submission Rules:

1. You must name your submission as specified in each exercise.
2. Upload the submission file on e-Centennial using the Assignment link(s).
3. In total you should submit the following:
 - a. One demonstration video, name it **firstname_comp264_assignment#2_demo**.
 - b. One analysis report for the exercise, the format is either as a word document or pdf, name it **firstname_comp264_assignment#2_analysis_report**. The report should have an intro and your name and student id at contain
 - c. One html file named as indicated in the exercise requirement.

Rubric

Evaluation criteria	Not acceptable	Below Average	Average	Competent	Excellent
	0% - 24%	25%-49%	50-69%	70%-83%	84%-100%
Requirements in exercise 40%	Missing all requirements required	Some requirements are implemented.	Majority of requirements are implemented but some are malfunctioning.	Majority of requirements implemented.	All requirements are implemented Correctly.
Written analysis report content 30%	Missed all the key ideas; very shallow.	Shows some thinking and reasoning but most ideas are underdeveloped. A few screenshots presented and some are partial.	Indicates thinking and reasoning applied with original thought on a few ideas. Majority of screenshots presented.	Indicates original thinking and develops ideas with sufficient and firm evidence. All screenshots are in order and fully presented.	Indicates synthesis of ideas, in-depth analysis and evidences original thought and support for the topic. All screenshots are in order and fully presented.
Written analysis report format and organization 5%	Writing lacks logical organization. It shows no coherence and ideas lack unity. Serious errors. No transitions. Format is very messy.	Writing lacks logical organization. It shows some coherence but ideas lack unity. Serious errors. Format needs attention, some major errors.	Writing is coherent and logically organized. Some points remain misplaced. Format is neat but has some assembly errors.	Writing is coherent and logically organized with transitions used between ideas and paragraphs to create coherence. Overall unity of ideas is present. Format is neat and correctly assembled.	Writing shows high degree of attention to logic and reasoning of all points. Unity clearly leads the reader to the conclusion. Format is neat and correctly assembled with professional look.
Demonstration Video 25%	Very weak no mention of the code changes. Execution of code not demonstrated.	Some parts of the code and code changes presented. Execution of code partially demonstrated.	All code and code changes presented but without explanation why. Code demonstrated.	All code and code changes presented with explanation, exceeding time limit. Code demonstrated.	A comprehensive view of code and all code changes presented with explanation, within time limit. Code demonstrated.

Demonstration Video Recording

Please record a short video (max 8 minutes) to explain/demonstrate your assignment solution. You may use the Windows 10 Game bar to do the recording:

1. Press the Windows key + G at the same time to open the Game Bar dialog.
2. Check the "Yes, this is a game" checkbox to load the Game Bar.
3. Click on the Start Recording button (or Win + Alt + R) to begin capturing the video.
4. Stop the recording by clicking on the red recording bar that will be on the top right of the program window.

(If it disappears on you, press Win + G again to bring the Game Bar back.)

You'll find your recorded video (MP4 file), under the Videos folder in a subfolder called Captures.

Or

You can use any other video recording package freely available