Assignment 1

1. PART I: Confirm UNT Student Email Address (10 Points)

Before answering the following questions, the student is required to read the document **VERIFY_UNT_email_address.pdf** posted in the Canvas module: **…/WEEK_1**.

The student is required to provide answers to the following questions:

1.1 What is the student's UNT student email address (**…@my.unt.edu**)?

1.2 Has the student received the Welcome-to-the-Class message via his/her email?

1.3 If NO to Question 1.2, can the student access the student email (**…@my.unt.edu**)?

1.4 Is the student a UNT employee?

1.5 If YES to Question 1.4, does he/she prefer using the UNT employee email for class communication?

1.6 If YES to Question 1.5, what is the student's UNT employee email?

IMPORTANT NOTES:

*--) We will use the **UNT email (…@my.unt.edu)** extensively in our communication. Please access your UNT email daily so that you can follow the class activities closely.*

--) If you cannot access your UNT email account for some reason, please let the instructor know as soon as possible.

2. PART II: Select an Operating System (10 Points)

During the course, the student uses cloud technologies in classwork. The student needs to use his/her local computing devices to access the remote deep learning server set up in cloud. For this purpose, it is required that the student should know how the operating system installed in his/her local computers work and can use it effectively.

TO-DO
--) **Select an operating system (OS)**, **either Windows, MAC, or Linux,** that **the student knows it well** and **can use it effectively.**

--) For the **selected operating system**, it is **required** that:
- The student should have fundamental knowledge about the operating system and its major, important components such as **drives, directories and folders, files, the ownership status of these components**, **the administrative privileges**, to name a few.
  - Fundamental knowledge includes, but not limited to:
    - How to **find where** a file or media contents are saved in the computing device
    - How to **get the real, physical path** of a file or a directory created or saved in the local storage of the computing device?
    - How to **access real, physical media contents** saved in the computing device
    - …

- The student should be able to run basic operations successfully on any computing devices in which the selected OS is installed.
  - Basic operations include, but not limited to:
    - **Creating new directory and folders**
    - **Accessing real, physical directories and folders and their contents**
    - Creating and saving new contents in the folders and can **access their real, physical files**, not only the alias names or links as in the MAC OS.
    - **Download** software applications and **save** them in a specified folder
    - **Install** software applications in a specified **real, physical** directory or folder
    - **Open terminals** in the windows environment of the OS
    - **Run command lines** in terminals in the windows environment of the OS
    - …

- The student should be **able to resolve issues** related to the **OS** and the **host computing device** if they happen while he/she uses the selected OS and its host computing device.
  - The student should **know how to do research**, e.g., using Google search, and **get technical support** such as contacting the vendors of the devices and the maker of the OS, and other supporting facilities like public technical forums, …

SUBMISSION REQUIREMENT PART II:
--) Provide a **short paragraph** to **specify** which operating system has been selected and to **state** that the student can use it effectively.

3. PART III: Set Up Deep Learning Virtual Machine (VM) in GCP (20 Points)

3.1 Overview: Cloud Computing and Google Cloud Platform (GCP)

The rise of cloud computing has been a facilitator to the emergence of big data. Cloud computing is the commodification of computing time and data storage using standardized technologies.

Big data is a term to describe large volumes of data that can be both structured and unstructured. These enormous volumes of data overwhelm the digital world every second. However, it is not the amount of data that is important. It is what we can do with the data that matters: Big data analytics can provide insights that lead to better decisions and strategic moves.

It was the emergence of cloud computing, which made it easier to provide the best of technology in the most cost-effective packages. Cloud computing has not only reduced costs but also made a wide array of applications available to companies of all sizes: small, mid-sized, big, and giant corporations.

Google Cloud Platform (GCP) provides on-demand cloud computing platforms to individuals, companies, and governments. The technology allows subscribers to have a full-fledged virtual cluster of computers, available all the time, through the Internet. Google Cloud Platform (GCP) virtual machines (VM) have most of the attributes of a real computer, including hardware (CPU(s) & GPU(s) for processing, local/RAM memory, hard-disk/SSD storage); a choice of operating systems; networking; and pre-loaded application software.

Each GCP virtual machine also virtualizes its console I/O (keyboard, display, and mouse), allowing the user to connect to their remote instances using a modern browser. The browser acts as a window into the virtual computer, letting subscribers log-in, configure and use their virtual systems just as they would a real physical computer.

3.2 Set Up Remote Deep Learning Server in GCP

IMPORTANT NOTES:

--) If the student already used the GCP free credit before, it is likely that he/she does not get the free credit again. In this case, the student must pay for using the GCP service.

TO-DO
--) Based on the lectures, set up a deep learning virtual machine (VM) in Google Cloud Platform (GCP).
--) Based on the lectures, install the GCLOUD SDK tool on the student's local computer.

SUBMISSION REQUIREMENT PART III #1:
--) Write a brief report to summarize the major steps of setting up the remote server.
--) Capture the screenshot that shows the critical information of the newly created remote instance.
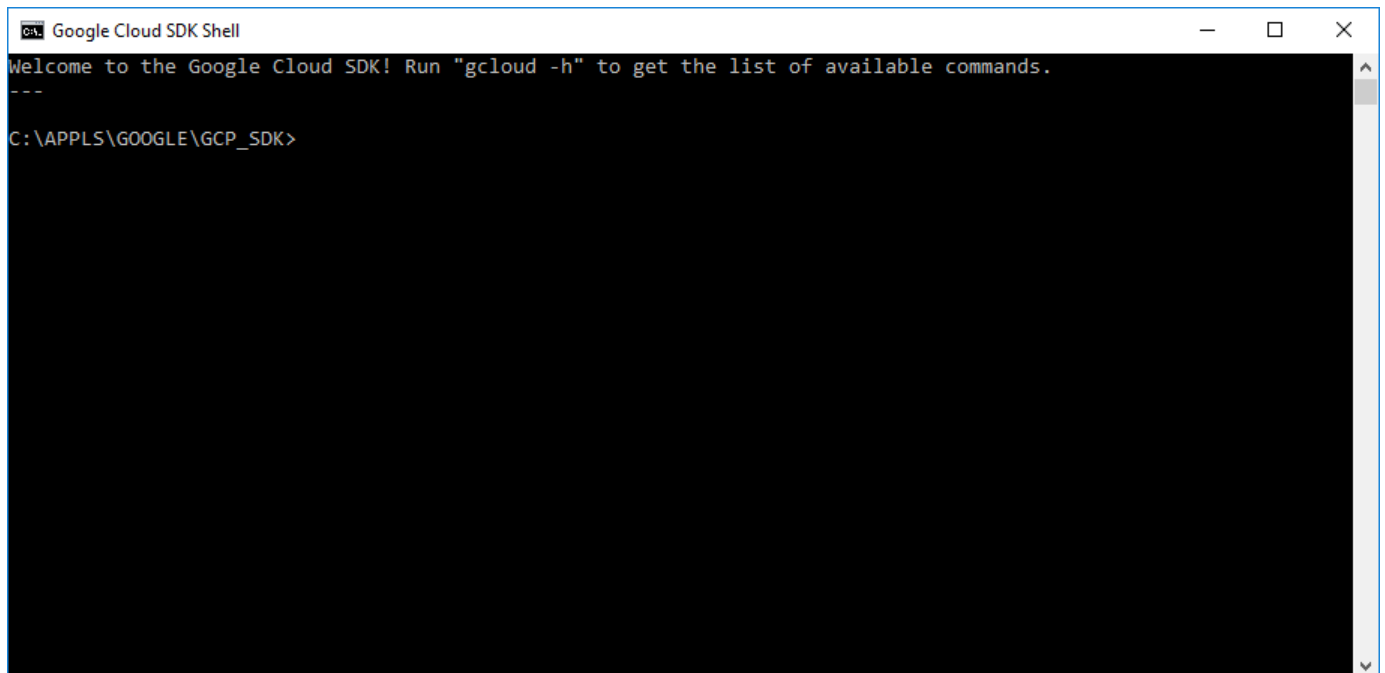
Here is one example of the screenshot:
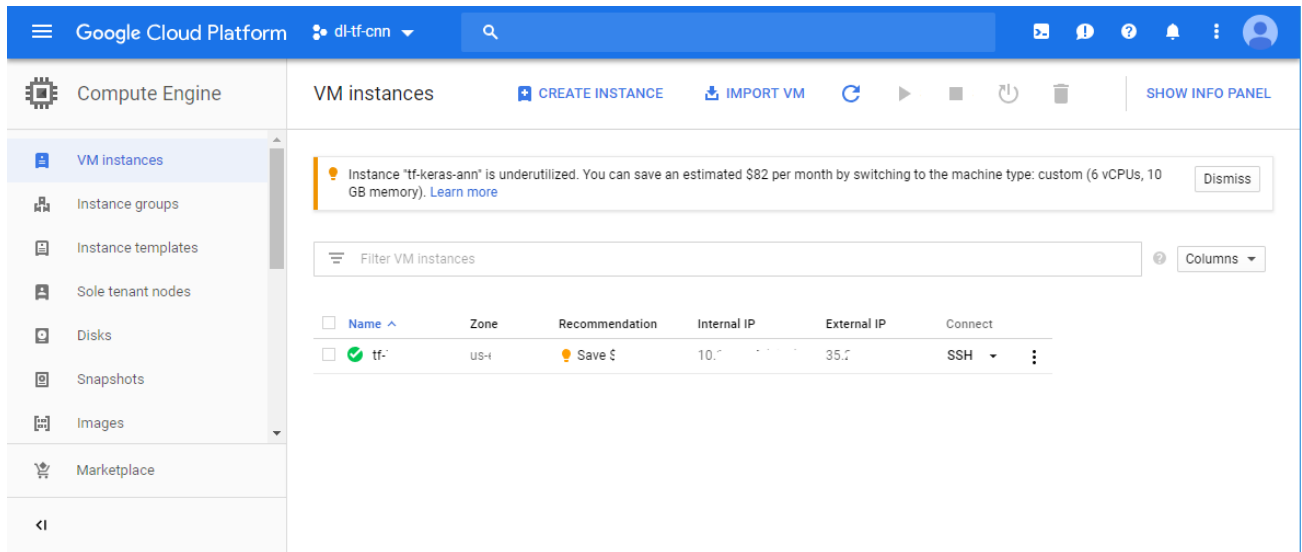
SUBMISSION REQUIREMENT PART III #2:
--) Write a brief report to summarize the major steps of installing the GCLOUD SDK.
--) Capture the screenshot that shows the tool has been successfully installed.

Here is one example of the screenshot:

4. PART IV: Connect Explore Remote VM Using SSH (10 Points)

TO-DO
--) **Question 4.1**:
Based on the lectures, open an SSH connection from the local computer to the remote VM.

--) **Question 4.2:**
Using the basic Linux command lines to explore the contents of the home directory

--) **Question 4.3**:
Create a new sub-folder named "JPTR_NTBK" under the home directory

--) **Question 4.4:**
Change the current directory to the newly created folder

SUBMISSION REQUIREMENT PART IV:
--) Write a brief report to describe the major activities the student has finished in PART IV
--) For each question, capture the screenshot that shows what he/she has done

IMPORTANT NOTES:
--) The answer should be clearly labeled with which question that the student is working on.

5. PART V: Start and Connect to Jupyter Notebook in Remote VM (30 Points)

TO-DO
--) **Question 5.1**:
Based on the lectures, **start** the Jupyter Notebook server in the **remote** virtual machine.

--) **Question 5.2**:
Connect to the **Jupyter Notebook** server in the **remote** virtual machine (by connecting a Local Computer Port, i.e., 8000, to the Remote Server Port, i.e., 8888)

--) **Question 5.3**:
Access and open Jupyter Notebook that is **currently running** in the **Remote Server** (in a browser on the local computer)

SUBMISSION REQUIREMENT PART V:
--) Write a brief report to explain the steps the student has done to start, connect, and use Jupyter Notebook that runs in the remote virtual machine.
--) For each question, capture the screenshot that shows what he/she has done

IMPORTANT NOTES:
--) The answer should be clearly labeled with which question that the student is working on.
*--) **Question 5.1, 5.2, 5.3**: The question does **not** ask to **start** Jupyter Notebook on the **local computer**. For example, if the student has installed an Anaconda package on his/her local computer. He/she can start Jupyter Notebook on his/her local computer. The question does **not** ask to do that.*

6. PART VI: History of AI and Big Data (20 Points)

Discuss (including **images** for illustration) the history of Artificial Intelligence from the '40s until now.

SUBMISSION REQUIREMENT PART VI:

It is expected that the student's work should include the following major sections:
--) The start of AI
--) The periods in which AI seemed to be stalled for many years and WHY?
--) The recent significant achievements, especially in the deep learning area, and the role of Big Data.

7. HOWTO Submit

The student is required to submit all the sections, i.e., submission requirements, in a Microsoft Word document that is sent to the instructor (Thuan.Nguyen@unt.edu) as an attachment to a UNT email.

The subject of the email must be: "ADTA 5560: Assignment 1 – Submission."

Due date & time: 11:00 PM – Wednesday 10/23/2024