

# MNIST notebook examples on IBM Watson Studio, Google VertexAI & Colab, and Amazon Sagemaker

NYU Cloud and machine Learning  
Spring 2023  
I-HSIN CHUNG and HAO YU  
Homework 2  
Xiaohan Wu (xw2788)

## Introduction

In this homework, I ran different MNIST notebook examples on IBM Watson Studio, Google VertexAI & Colab, and Amazon Sagemaker, which are machine learning platform services provided by IBM cloud, Google cloud, and AWS. The notebook examples use storage and computing resources on the cloud to deploy models to predict MNIST handwritten digits. A comparison is made after successfully running the code on the clouds.

## My works on clouds

### IBM Watson Studio

With the help of a clear instruction on hands-on slides, it is fast to accomplish the task. All the resources are created when we created the project and we can configure that later. The example notebook is also selected during the creation process.

### Amazon Sagemaker

It is easy to create a notebook instance in Sagemaker and work on the task. But it took me some time to create an executable account. I assume that professors have carefully set up our accounts on IBM so we don't need to worry about that. But on AWS, when I ran the code, some errors on resources limit, access limit, occurred and it turned out that it is because I didn't grant enough permissions to the executable account or the project. I didn't manage to find a good example for Sagemaker, so I asked professor Yu for help and he gave me an example in the example repository but not in the index. The code is fine as long as it has enough permissions and it saves and loads models with Amazon s3.

### Google VertexAI

Following the clear document from google, I quickly created a notebook to work on the task. The examples are directly given in the instance. But I spent a lot of time debugging the code because the MNIST examples use tf1 which is not compatible with tf2. The example code uses google storage(GS) to save models.

### Google Colab

I also included Colab code because it is also a popular platform for machine learning on google cloud. It is different from the previous three for it is stored in google drive. It is easy to create and easy to use just like working on google docs and google sheets. When I found this code, I gave it a go and it runs well.

## Comparison

### Common points

1. The documents are complete and up-to-date.
2. Have account permission management, including collaboration and restrictions.
3. Instances, storage, and spaces are provided by the platforms
4. All the example code is open-source and is available at Github
5. Provide CLI tools to connect to cloud to use service

### Example library

Watson Studio provides examples when creating a project, it will ask users whether to use an existing project or a template to start. Details for running the project are also included in the examples, encountering no questions when trying to run the code.

Google Vertex AI provides the most number of examples among the three examples and it is directly provided in the instance. Besides official examples, it also has a community folder which also contains excellent example code. But not all examples are up-to-date. For example, the MNIST examples I am running for the homework are written in tf1. Users who want to run that out-dated code, will need to either change the tf versions in the environment or translate their code. Different from opening a project from given options, Google Cloud supports open example code through a badge which is very flexible.

Amazon Sagemaker has a repository of examples like the other 2 platforms. But Users will have to get it from Github. No quick access is provided.

### Connect to notebook from laptop

All of the three cloud services provide CLI tools, users can use terminals to manage their clouds without using a Web based UI. However, as for using local notebook to connect cloud notebook, Watson Studio gives the best experience, it provides the API to connect to the cloud project to run the application on the cloud. Google Cloud recommends using their services on the web-based editor and Amazon Sagemaker provides a way to run a notebook on the cloud then configure the local environment to connect to the cloud notebook, which is not as easy and fast as Watson Studio.

## References

1. "IBM Cloud Virtual Servers" [online]. Available: <https://www.ibm.com/cloud/virtual-servers>
2. "Machine Learning - Amazon Sagemaker" [online]. Available: <https://aws.amazon.com/sagemaker/>
3. "Amazon Simple Storage Service (Amazon S3)". [online]. Available: <https://aws.amazon.com/s3/>
4. "Vertex AI | Google Cloud" [online]. Available: <https://cloud.google.com/vertex-ai>
5. "Cloud Storage | Google Cloud" [online]. Available: <https://cloud.google.com/storage>
6. "Google Colab" [online]. Available: <https://colab.research.google.com/>