**Activity 8 : Machine Learning**

| Group # |  |
| --- | --- |

Participating group members

| Student ID | Name |
| --- | --- |
|  |  |
|  |  |
|  |  |

Outline

Part A: Feature Engineering

Goal: 1. Practice feature engineering -- deriving predictive features from data.  
2. Understand the basic process of building a machine learning model.  
3. Establish basic knowledge of data sets through exploratory data analysis.

Part B: Image Classification

Goal: 1. Practice data gathering skills.  
2. Understand transfer learning.  
3. See the state-of-the-art techniques in ML/AI in actions.

Outstanding Factor

1. Accuracy of Part A on test set is more than 73%

2. Model of Part B classifies all test images correctly.

# 

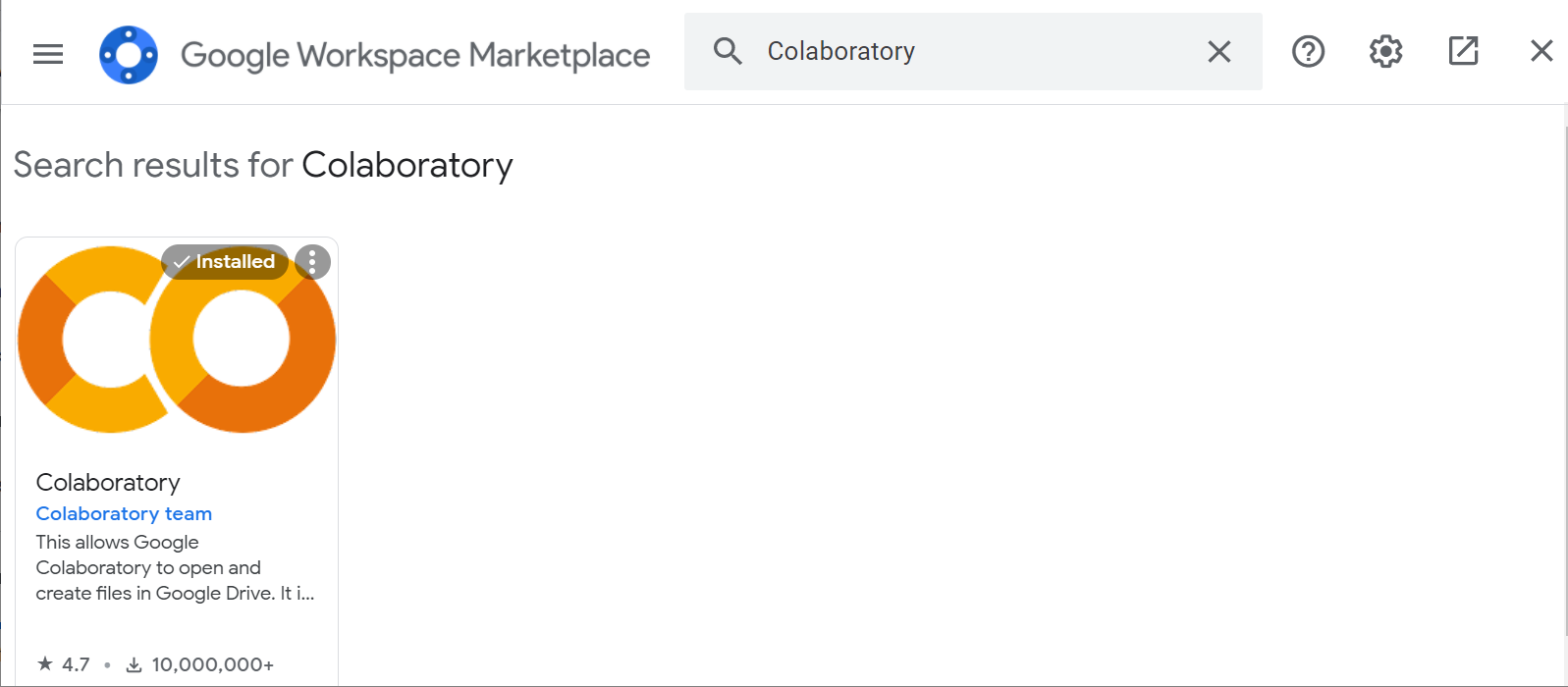
Part A: Feature Engineering

Understand the process of building a machine learning model

## Instruction

1. Download extracted zip file from “Activity 8” material in CourseVille.
2. Copy “**activity8.1-feature-engineering.ipynb**” to your Google Drive.
3. Install Google Colaboratory on your Google Drive.

* Right click on the file >>> Open with >>> Connect more apps
* Search for “Colaboratory”
* Install



1. Open the .ipynb file with Google Colab.

* Right click on the file >>> Open with >>> Google Colaboratory

1. Edit and Run Code to answer Question 1 to 7.

## Questions

1. Given a post with 500 shares, which page is it more likely to come from?

|  |
| --- |

1. Given a post published at 20:30, which page is it more likely to come from?

|  |
| --- |

1. Given a post with 140 characters, which page is it more likely to come from?

|  |
| --- |

1. Discuss how you come up with the list of words.

|  |
| --- |

1. Write down your final list of words together with their feature importances and the training accuracy. Also include a screenshot of the training result.

|  |
| --- |

1. Write down your test accuracy, also include a screenshot of the testing result. Discuss how it differs from the training accuracy.

|  |
| --- |

1. There will be the “tree.svg” File console of Google Colab. Upload it to Google Drive and put the link here.

|  |
| --- |

# 

Part B: Image Classification  
See the Deep Learning techniques in ML/AI in actions

## Instruction

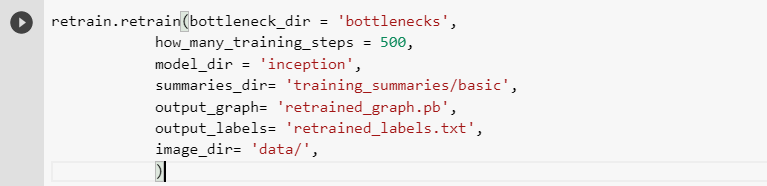
1. Brings items and video clips as prepared in pre-test.
2. Open the “**activity8.2-Image-Classification.ipynb**” file with Google Colab.
3. Install Tensorflow version 1.15
4. Sample images from those videos using [ffmpeg](https://www.ffmpeg.org/download.html).

ffmpeg -i myvideo.mov -vf fps=1 img%03d.jpg

This will output one image every second, named img001.jpg, img002.jpg, img003.jpg, etc. You can generate more photos by adjusting the fps to 2 or 3. The %03d dictates that the ordinal number of each output image will be formatted using 3 digits.  
 should be put in separate folders accImagesordingly:



1. Create your own classifier by run this cell.



1. While waiting for the retraining process, let’s explore Google’s Cloud Vision API (<https://cloud.google.com/vision/>). Pick images from each class and input them to the demo. Discuss what Google sees in those images.
2. Test your model with images in the “Test Data” folder! Can they distinguish your 5 items? **(Ask password for test.rar from TA)**

## Questions

1. Label all test images one by one. Put screenshots of your results along with corresponding test images filename below.  
    Answering Example:

| 1. pillow1.jpg |
| --- |

|  |
| --- |

1. Zip all of your images and upload to google drive (with identical folder structure) and put the link below.  
   (name your file like this “<GroupName>\_activity8.zip”)

|  |
| --- |

# 

Outstanding Factor

Accuracy of Part A on test set is more than 73%

Model of Part B classifies all test images correctly.