**Setup Instructions for Automating Machine Learning with Paperspace**

# Purpose

The purpose of this application is to automate the setup and use of Machine Learning through Paperspace. This is a great way to get started using Machine Learning, as all the configuration is done for you and you can get right into working with the notebook.

Firstly, the application automates account setup, meaning you can enter in your desired email/password combination and the script will use these credentials to set up an account. It will then prompt you to check your email and confirm the account. It is now ready to use. After you tell the application that you confirmed the email, it will generate an API key and use it to link your computer and paperspace account.

An area where first time users may get confused is configuring the notebook. There are an abundance of different containers and libraries to use, but the easiest and most robust for general use is fast.ai. There are lots of useful lessons which can be found here ( <https://course19.fast.ai> ) – from here, you can choose if you would like to build a model which classifies images, or which processes language. The tutorials are simple and easy to understand, but you can also just get one from the internet if you want.

The main benefit of using this application is the fact you can always use a free GPU. This circumvents the need to use your own PC, or to buy GPU usage on the site. The application will attempt to start the notebook with a Free GPU, and if one is not available, it will keep retrying. These free GPUs have an upper limit of 6 hours, meaning you can do 6 hours of free processing/training before it will shut off. This is plenty of time for an entry-level model to finish training. Also, this is extremely useful for multiple models training in parallel, for example. These are the installs necessary to get a program like this up and running.

# Potential Uses

There are many potential uses for this application. If you wanted to train a model overnight, you could start the application before leaving your PC and when you come back in the morning you will have a trained model ready for use. This gets rid of the need to keep waiting for a free GPU to become available, for the dataset to upload, or for your model to finish training – saving a lot of time.

Also, this could potentially be used to train similar models in parallel using different paperspace accounts. Through this application, it is easy to set up an account just through entering your email. If you needed to, you could deploy multiple models with slightly different datasets to see how the results differ – or, maybe you just need to train multiple models at once but don’t have the hardware capabilities to do so.

# Necessary Installs

**Fast AI and Image Classification**

Fast AI Image Classification - <https://course.fast.ai/videos/?lesson=1>

Using Paperspace Gradient with Fast AI - <https://course.fast.ai/start_gradient.html>

**Python Applications**

VSCode – <https://code.visualstudio.com/download>

Install Python3 on Machine - <https://www.python.org/downloads/windows/>

Getting started with Python on VSCode - <https://code.visualstudio.com/docs/python/python-tutorial>

Add Python to PATH in Environment Variables - <https://geek-university.com/python/add-python-to-the-windows-path/>

Chromedriver – Used with Selenium to launch chrome <https://chromedriver.chromium.org/downloads>

PIP – Used to Install other packages - <https://phoenixnap.com/kb/install-pip-windows>

After installing and configuring Python and VSCode, certain packages are needed for use with it. The Paperspace package is vital and allows us to communicate with the Paperspace servers directly from the PC, with an API Key. Selenium is the tool used for automation, and PySimpleGUI is what is used to present the script in an easy to use GUI.

**Packages (pip install XXXXX)**

Paperspace – Used to interface with ML Apps hosted on Paperspace - <https://pypi.org/project/paperspace/>

Selenium – Used to navigate through webpages e.g. for account setup - <https://selenium-python.readthedocs.io/getting-started.html>

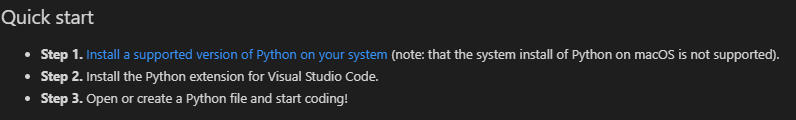
PySimpleGUI – An easy GUI so you can interact with the program without the command line - <https://pypi.org/project/PySimpleGUI/>

Also, this could potentially be used to train similar models in parallel using different paperspace accounts. Through this application, it is easy to set up an account just through entering your email. If you needed to, you could deploy multiple models with slightly different datasets to see how the results differ – or, maybe you just need to train multiple models at once but don’t have the hardware capabilities to do so.

# Setup Instructions

Firstly, you will need to download VSCode and configure it for working with Python. You can do this through downloading it [here](https://code.visualstudio.com/download).

After downloading it, you will need to make sure Python3 is installed on your machine. After installing this [here](https://www.python.org/downloads/windows/), you will need to install the extension on VSCode too.



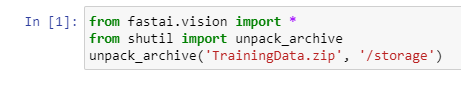
When this is done, it’s time to install the three packages. Go to the terminal via VSCode, and type in these instructions. If it doesn’t work, you may have an error because the path to the Python executable wasn’t added to the Windows Path variable. You can fix this [here](https://geek-university.com/python/add-python-to-the-windows-path/).







For this script to work, your notebook will have to be in a linear format. This means that all the buttons should be pressed in series, and there should be no jumping back and forth between cells. This layout is necessary for the program to work.



*from shutil import unpack\_archive*

*unpack\_archive('TrainingData.zip', '/storage')*

**- This code needs to go at the start, to unzip the dataset**

In the first cell of your notebook, make sure you add this code to unzip the training data to the storage folder. Inside the zip file, there should be a folder filled with images for each dataset. Make sure that this folder is formatted correctly, or the training won’t work!

# Useful Tools

<https://chrome.google.com/webstore/detail/download-all-images/ifipmflagepipjokmbdecpmjbibjnakm/related?hl=en> – Image Downloader for creating datasets