

Configuration Manual

Detection and Classification of Leaf Diseases in Maize Plant
using Machine Learning
MSc in Data Analytics

Adarsh Jayakumar
Student ID: x18131379

School of Computing
National College of Ireland

Supervisor: Dr. Cristina Muntean

National College of Ireland
MSc Project Submission Sheet
School of Computing



Student Name: Adarsh Jayakumar
Student ID: x18131379
Programme: MSc in Data Analytics **Year:** 2019-2020
Module: Research Project
Lecturer: Dr.Cristina Muntean
Submission Due Date: 12th December 2019
Project Title: Detection and Classification of Leaf Diseases in Maize Plant using Machine Learning
Word Count: 1000 **Page Count:** 6

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:
Date: 12th December 2019

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	<input type="checkbox"/>
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	<input type="checkbox"/>
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Adarsh Jayakumar
Student ID: 18131379

1 Introduction

This manual explains the steps and procedures that must be followed by any researcher, prior to running the scripts developed for the current research project. This will enable them to run the code smoothly without any setbacks. It also entails information on the hardware configuration of the machine in which the scripts were run and provides minimum recommended configuration for the same. Following these steps would help them replicate the results showcased by the project. This can then be analysed, and future research can be conducted easily.

2 System Specification

2.1 Hardware

The hardware configuration of the machine on which the research was conducted is as described below:

Processor: Intel i7 – 8750H CPU @ 2.20GHz

RAM: 16 GB

Storage: 256GB SSD + 1 TB

Operating System: 64-bit operating system, Windows 10 Home

The minimum configuration however can be much lower. The dataset used in the project has close to 4000 images of sizes in and around 300*300 pixels and this can be loaded and analysed using systems having around 8gb of ram and an i5 processor. The operating system can either be windows or Mac. However, one must note that lower the specification of hardware, longer the duration taken by models to train on the data which can be a severe drawback.

2.2 Software

The software chosen for executing the project is the Jupyter Notebook Integrated Development Environment (IDE) that is based on python and is available as part of

Anaconda package. The installations and steps to be followed in running the scripts developed are described in detail in the upcoming section.

3 Downloads and Installation

- **Python**

The research project is implemented using python as it has large number of libraries that support machine and deep learning models. It also comes with several modules that help in easier pre-processing and manipulation of images, thus enabling ease of use and implementation. So, the basic requirement on the machine that would run the script is to have the latest version of python installed on it. This can be done by heading over to the download section of python website¹ and downloading the executable installer of the version chosen based on OS of the machine running it. Figure 1 shows the screenshot of the page from where the latest version can be downloaded. Once, the executable has been downloaded, it is to be installed by following the instructions provided by the installer.

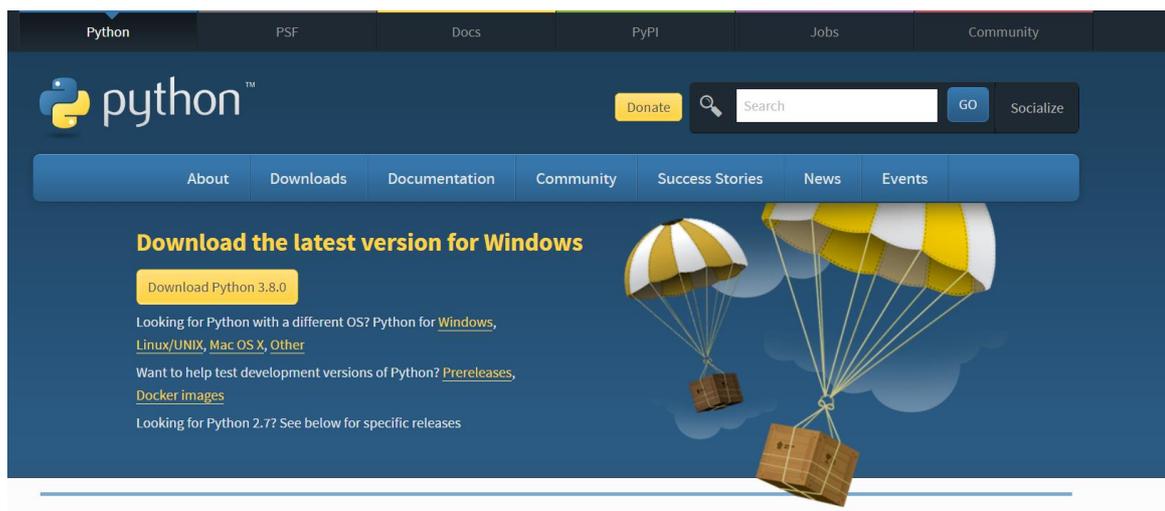


Figure 1. Python Download Page

The successful installation of the same can be verified using the ‘python –version’ command on windows command prompt terminal, which, if successful returns the version number of python installed.

¹ <https://www.python.org/downloads/>

- **Anaconda**

The next package that must be installed is Anaconda. It provides various IDE's based on python that are user friendly that can be used to develop the code and visualise the results. Of the IDE's available on installation, Jupyter Notebook and Spyder are the most popular. Anaconda² package can be downloaded from their home website and a screen shot of the page from which the installer is downloaded is shown in Figure 2. As mentioned in case of python, it also comes with different installers for different operating system and must be chosen based on machine running the script.

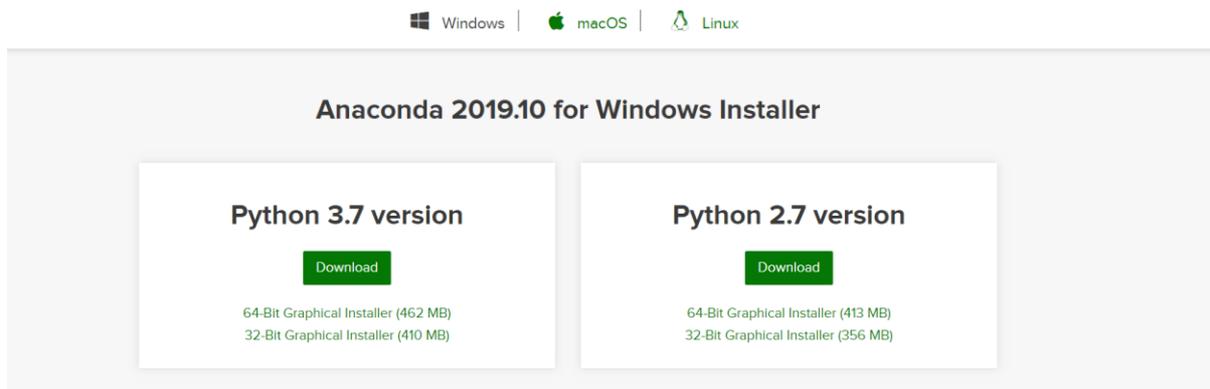


Figure 2. Anaconda Download Page

The Anaconda Navigator, upon successful installation will display a screen as shown in Figure 3 from which the desired IDE can be chosen for development. Jupyter Notebook has been used in the current research.

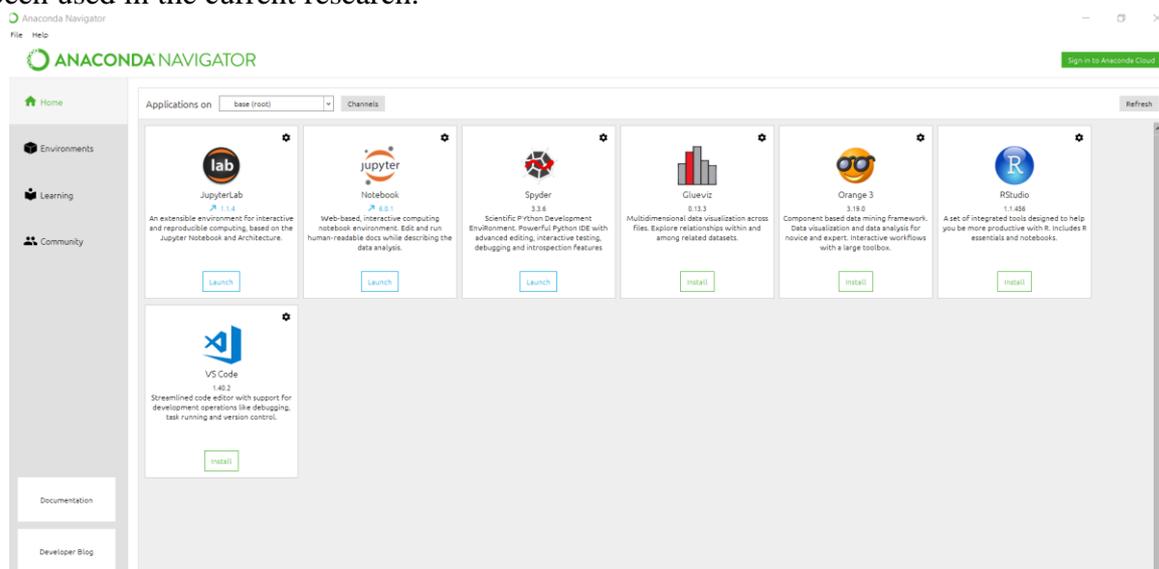


Figure 3. Anaconda Navigator

² <https://www.anaconda.com/distribution/>

- **Data source**

The dataset used for the project must be downloaded from GitHub³ and stored on the machine in which the research is conducted. The dataset has images of many plants (Figure 4) along with coloured and grey scale images of the same. The current research makes use of only coloured version of maize plant images and the remaining folders are deleted from the local database.

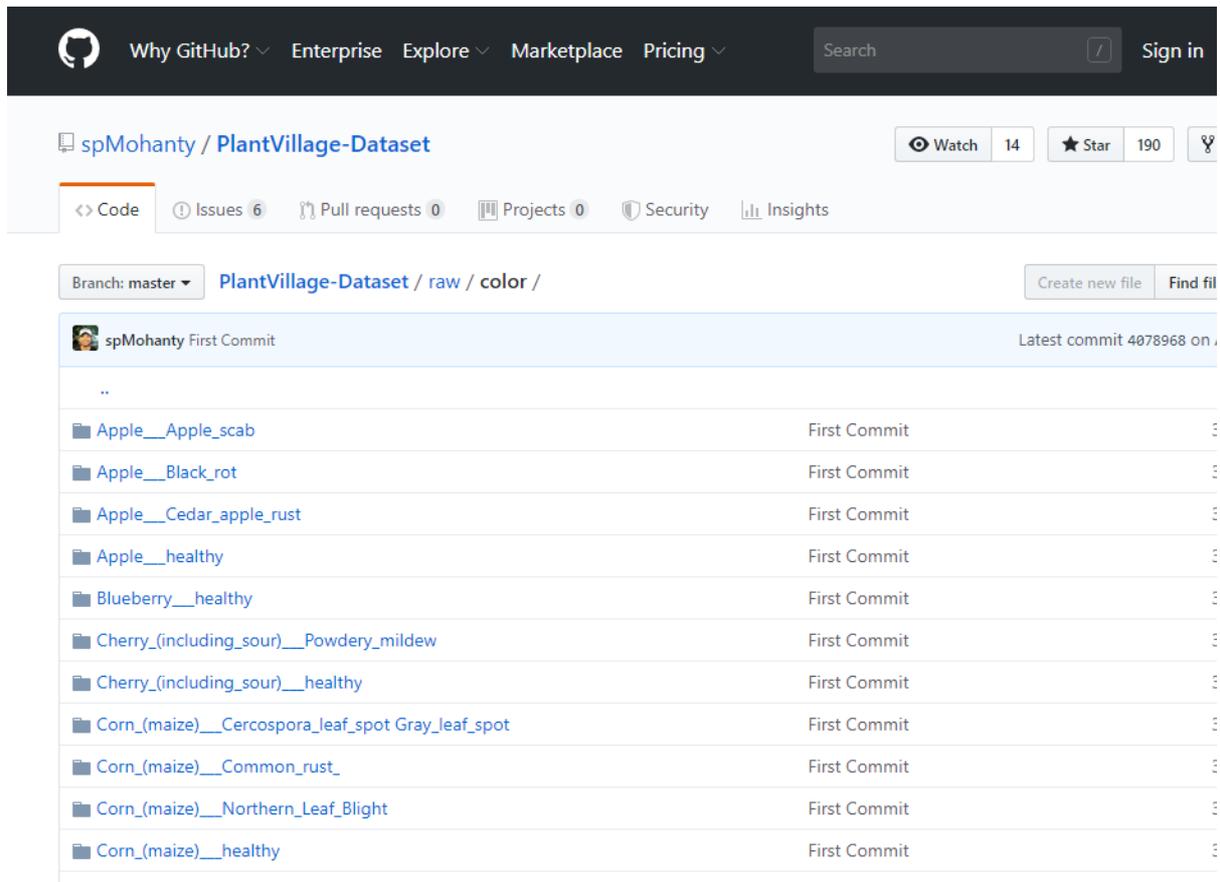


Figure 4. Dataset Used

4 Project Development

Jupyter notebook is to be launched from the navigator installed. The application opens as a tab in browser as shown in Figure 5.

³ <https://github.com/spMohanty/PlantVillage-Dataset>

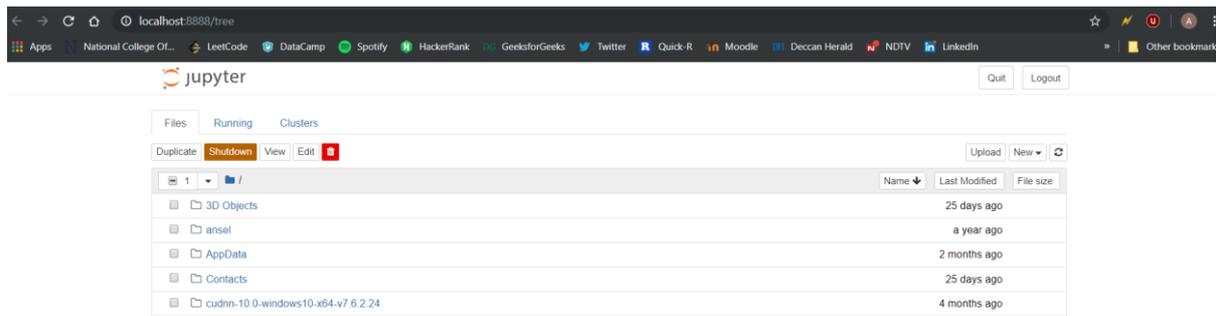


Figure 5. Jupyter Notebook Home Page

Clicking the new icon and selecting python 3 will launch a page where one can start developing the code. As the project is developed using machine and deep learning techniques based on transfer learning, one must also install python libraries required for their development as and when they are needed. However, some of the standard libraries that are a must for development of image classification models can be installed initially and these include

- TensorFlow 2.0.0
- Keras 2.3.1
- Keras-Applications 1.0.8
- Keras-Preprocessing 1.1.0
- Numpy 1.16.5
- Scikit-Image 0.16.2
- Scikit-Learn 0.21.3
- Sklearn 0.0
- Opencv-contrib-python 4.1.1.26
- Matplotlib 3.1.1

These are installed using the command prompt of windows by using the pip install command. For eg. 'pip install tensorflow' as shown in Figure 6.



Figure 6. Python library installations in command prompt

Upon successful completion of writing the script, one can execute the script using run command of jupyter notebook as shown in Figure 7 or by navigating to cells and using run all

option to execute all the code blocks. If errors are present, they will be displayed below the code block and can be used for debugging later.

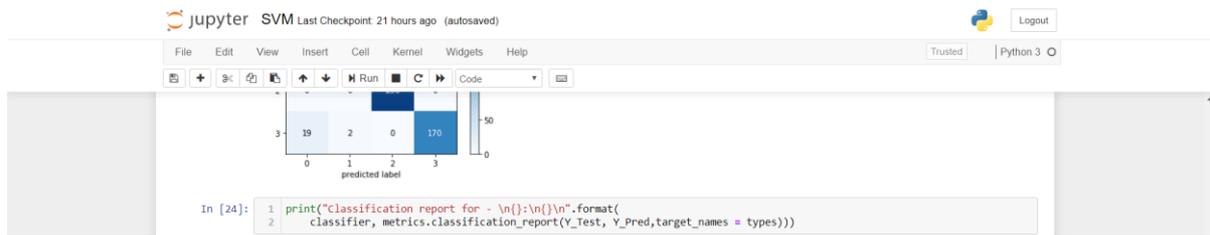


Figure 7. Running the script

If all the code blocks of the algorithm are run successfully, one can analyze the results obtained and compare it with other models to evaluate each of their efficiency in classifying and identifying maize disease. The results obtained for a successful execution of a code block for SVM model is as shown in Figure 8 and Figure 9.

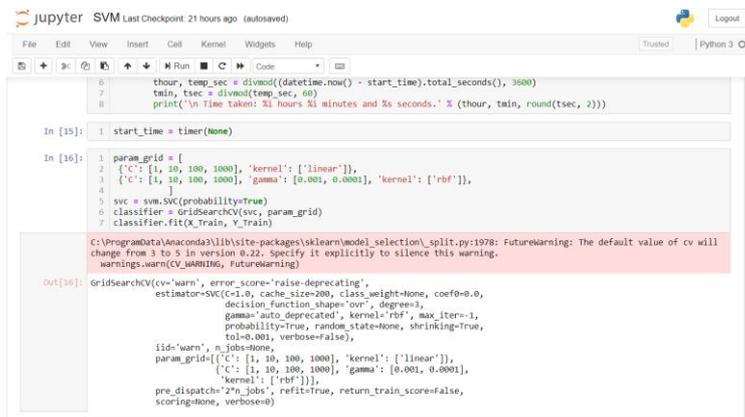


Figure 8. SVM Code for Classification

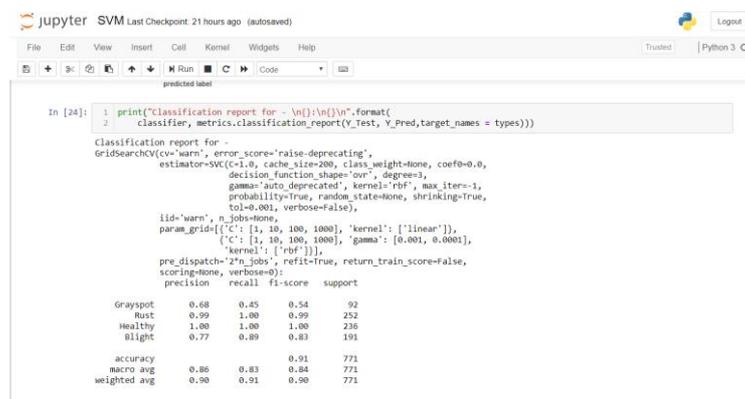


Figure 9. Classification Report for SVM

The scripts can then be downloaded as python or jupyter notebook. All the remaining scripts are submitted along with dataset as part of ICT solution