**DATA 255 - Group 2**

*Exploring the Potential of Deep Learning Architectures for Plant Trait Prediction*

**Dataset**

The dataset was taken from Kaggle which is now remove so adding the data in [google drive](https://drive.google.com/drive/folders/1eeQD8SnKG1AwECQ7IWY2f5JbWYtQpEg0?usp=sharing) and sharing link. The data is 3.17GB which has one directory named images and on csv file named tab\_data.csv. Images has names which is referenced in csv file as id. Download the data and place it in a directory named “data”. The structure should be as shown below once the data (images/\* and tab\_data.csv is downloaded into data).

DATA255\_Group2\_Project

|─ code

|   ├── buildModel.py

|   ├── loadData.py

|   ├── main.py

|   ├── preprocessData.py

|   ├── setupGPUS.py

|   └── visualization.py

|─ data

|   ├── images

|   └── tab\_data.csv

|─ project\_report

|   ├── DATA255\_Group2\_ProjectReport.pdf

|   └── DATA255\_Group2\_ProjectReport\_latex.tex

|─ readme.docx

|─ requirements.txt

**Steps to run the code:**

**Prerequisites:** Conda and Linux/MacOS

Once the data is downloaded and placed in the “data” folder,

1. Open the terminal in the unzipped folder and create a new environment using conda

`*conda create -n env\_name python*`

1. Activate the environment using

`*conda activate env\_name*`

1. Install all the libraries mentioned in requirements.txt from the conda environment

`*pip install -r requirements.txt*`

1. Once all the libraries are installed go into code using

`*cd code*`

1. And run

`*python main.py*`

Following the above steps will run the main.py and modelling will start and run for 50 epochs giving the results and showing loss curve with evaluation metric. Running this in HPC lab each epoch took 40s so for the program to run in whole it would take ~40mins including load, preprocess and modelling.

**Team contribution table**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Task** | **Bhavika** | **Gayatri** | **Sai Srivathsav** | **Shashank** |
| Background research | Yes | Yes | Yes | Yes |
| Literature review | Yes | Yes | Yes | Yes |
| Data Collection | Yes | Yes | Yes | Yes |
| Data Processing | Yes | Yes | Yes | Yes |
| Modeling | MobileNetV2, TabNet | EfficientNetB0 and B2 | MobileNetV2 + DenseFC and ResNet18 | EfficientNetV2 and VGG + LSTM |
| Documentation | Yes | Yes | Yes | Yes |