



Python and Deep Learning Programming

CSEE 5590 0001

Spring 2020

Project Plan Report

Team Members:

Sai Harshavardhan Maddula(class Id: 15)

Venkata Sri Rohita Goparaju(class Id: 27)

Divya Reddy Bandari(class Id: 7)

Sai Srinivas Vidiyala(class Id: 17)

Plant Seedlings Classification

(Determine the species of a seedling from an image)

Kaggle Link:

The data from Kaggle is Plant-Seedlings Classification

<https://www.kaggle.com/c/plant-seedlings-classification/overview>

Inspiration:

Me and my teammates found this project interesting for of the following reason, since it focuses on:

- Differentiating a weed from a crop seedling. Is that possible?
- Classifying the data through images by training with images.

Project Overview:

Achieving the agenda will give a better yield on crops and better stewardship of the environment.

Project Description:

The Aarhus University Signal Processing group, in collaboration with University of Southern Denmark, has recently released a dataset containing images of approximately 960 unique plants belonging to 12 species at several growth stages.



The dataset is also hosted as a Kaggle competition in order to have a wider exposure, to give the community an opportunity to experiment with different image recognition techniques, as well to provide a place to cross-pollinate ideas.

Data Description:

The dataset has about 4750 labeled images (1.73GB) showing plants of 12 different types. The goal is to classify correctly the species shown on the 794 images (91MB) of the test set. All images are quadratic but vary in size. We are provided with a training set and a test set of images of plant seedlings at various stages of grown. Each image has a filename that is its unique id. The dataset comprises 12 plant species. The goal of the competition is to create a classifier capable of determining a plant's species from a photo. The list of species is as follows:

- Black-grass
- Charlock
- Cleavers
- Common Chickweed
- Common wheat
- Fat Hen
- Loose Silky-bent
- Maize
- Scentless Mayweed
- Shepherds Purse
- Small-flowered Cranesbill
- Sugar beet

Dataset Description:

- train.csv - the training set, with plant species organized by folder
- test.csv - the test set, you need to predict the species of each image
- sample_submission.csv - a sample submission file in the correct format

Conclusion:

- Using a subset of the samples provided in the dataset of each species for training, we can achieve an accuracy.
- Using all labeled data for training and making use of little data augmentation, we can obtain an more accuracy on the test set.