**Project Overview**

As University of Pretoria stated, a successful cultivation of crops broadly depends on the weed control strategies. It has been widely observed that tons of cultivated crops are wasted due to crop infestations. A zero percent loss in crops is entirely a rare event, generally there is 10 to 100% losses as per the current weed control practices. (Pretoria, n.d.)

Additionally, for farmers it is of vital importance to detect weed in the initial first to six weeks of plantation, because both weed and crop vigorously search of nutrients and water in the soil for this specific period.

According to the research paper published by Thomas Mosgaard Giselsson and co, there is no robust computer vision system out there, which can classify the species of ground-based weed. (Giselsson, Jørgensen, Jensen, Dyrmann, & Midtiby, 2017)

In this project, I have developed a CNN model which can detect plant species up to 92% of accuracy.

**Problem Statement**

Here, I am using a dataset which contains 5,539 images of crop and weed seedlings. The images are grouped in 12 classes. Moreover, these classes represent common plant species from Danish Agriculture. (Dyrmann)

Each classes contain RGB images, which reflect various stages of plant growth.

Using this dataset, the goal is to build a model which can further classify weed seedlings and crops.

Moreover, I have also planned to further extend this project by integrating this Deep Learning API with a web or mobile application, where a farmer will just have to upload the image in the mobile or web application, and the API will be able to predict the species of the weed or crop, thus, farmer can take appropriate action. In this way, if a farmer found a specific seedling as a weed, it can be destroyed before it infestate the actual crop.

The actual data set can be found out from the link mentioned below.

<https://vision.eng.au.dk/plant-seedlings-dataset/>

**Metrics**

The development of machine learning models revolves around the idea of constructive feedback principle. When we build a model, the evaluation metrics are used to get feedbacks, and thus further improvement is continued to achieve a desired accuracy.

Simply building a predictive model is not the prime motive, we must utilize the capability of evaluation metrices to discriminate among model results.

In this project, the evaluation metrices that I have used to validate the model, and for further improvement of model accuracy are Validation Accuracy, Precision, Recall, F-1 score, and Confusion Matrix.

* Validation Accuracy is where both true positive and true negative is taking into consideration.

Validation Accuracy =

Chart, line chart

Description automatically generated

* Precision: It tells the accuracy of positive predictions.

*Precision =*

* Recall defines the fraction of positive cases that the model is able to catch or correctly identified.

*Recall =*

* F1 scores tells the percent of correct positive predictions. It is nothing but the weighted harmonic mean of precision and recall such that the best score is 1.0 and the worst score is 0.

*F1 Score =*

A picture containing text, black

Description automatically generated