Group Lab 3

End to End Problem Solving
Increase Agricultural Productivity Using Machine Learning

Learning Objective:

At the end of this group lab experiment, you will have a complete end-to-end understanding of Machine Learning application in a specific domain i.e. Agriculture.

Domain Background:

Research shows that the impact of AI/ML in Agriculture is USD 150 Billion. Hence the study of Deep Learning in agriculture is very exciting.

Artificial intelligence holds the promise of driving an agricultural revolution at a time when the world must produce more food using fewer resources. Key Questions are: How are farms using artificial intelligence to direct crop planting, harvesting and more, and how are they getting that data in the first place? Can we use AI in the agriculture supply chain?

Experiment:

For this group lab we expect you to pick or formulate any **one of the problems** that is related to Agriculture. Please read Deep Learning in Agriculture: A Survey to understand the related domain in detail.

Here is a simple framework that might help you build an end to end problem.

Would you be able to prioritize the **problem** you have considered above, based on the following (but not limited to):

- Impact of the solution
- Ease of finding the solution
- Availability of support infrastructure
- Cost of implementation
- Acceptability of the solution in the society
- Feasibility of deployment of the solution
- Necessary skill level required to use the solution

Questionnaire to be answered in the Presentation:

Your presentation must answer the below questions for the problem that you have formulated:

a. What are the five most important pieces of information you need to solve the problem (chosen) using Machine learning? Why? What are the sources for that information?
b. Is the final problem supervised, semi supervised or unsupervised? Describe.
c. What are the deficiencies that you are likely to encounter with the data collected?

d. How do you preprocess/represent the data, in order to feed it to the algorithm?
e. What is the algorithm that you choose for solving the problem with this data in this setting? Why?
f. What is the rubric you would choose to evaluate and why?
g. How do know that the solution obtained is useful?
h. Brief the plan for entire process of model building and experimentation with all relevant details.

Evaluation Criterion and Grading Scheme: Total Marks: 20

Prepare a 10 slide presentation using the template.

Each Team will have 8 minutes to present in their respective labs.

The Rubric of evaluation is the following:

a) Problem Statement (5M)
b) Responses to the Questionnaire and Concepts (5M)
c) Overall Presentation (10M)

Total Marks = 50% of marks awarded by lab mentor out of 20 + 50% of average of the marks awarded by the rest of the teams(one SPOC only) in the lab.

Example: Lab Mentor: 16 (50 % is 8 Marks) ; Team 1: 18; Team 2: 16; Team 3: 14

Average of Team 1,2 and 3 = (18 +16+14)/3 = 16 (50% of 16 is 8 Marks)

Total Marks: 8+8 = 16 Marks

Quick Reference to Datasets:

- You are free to use any other datasets as well.

No.	Organization/Dataset	Description of dataset	Source
1	Image-Net Dataset	Images of various plants (trees, vegetables, flowers)	http://image-net.org/explore?wnid=n07707451
2	ImageNet Large Scale Visual Recognition Challenge (ILSVRC)	Images that allow object localization and detection	http://image-net.org/challenges/LSVRC/2017/#det
3	University of Arkansas, Plants Dataset	Herbicide injury image database	https://plants.uaex.edu/herbicide/ http://www.uaex.edu/yard-garden/resource-library/diseases/

4	EPFL, Plant Village Dataset	Images of various crops and their diseases	https://www.plantvillage.org/en/crops
5	Leafsnap Dataset	Leaves from 185 tree species from the Northeastern United States.	http://leafsnap.com/dataset/
6	LifeCLEF Dataset	Identity, geographic distribution and uses of plants	http://www.imageclef.org/2014/lifeclef/plant
7	PASCAL Visual Object Classes Dataset	Images of various animals (birds, cats,horses, sheep etc. cows, dogs,	http://host.robots.ox.ac.uk/pascal/VOC/
8	Africa Soil Information Service (AFSIS) dataset	Continent-wide digital soil maps for Sub-Saharan Africa	http://africasoils.net/services/data/
9	UC Merced Land Use Dataset	A 21 class land use image dataset	http://vision.ucmerced.edu/datasets/landuse.html
10	MalayaKew Dataset	Scan-like images of leaves from 44 species classes.	http://web.fsktm.um.edu.my/~cschan/downloads_MKLeaf_dataset.html
11	Crop/Weed Field Image Dataset	Field images, vegetation segmentation masks and crop/weed plant type annotations.	https://pdfs.semanticscholar.org/58a0/9b1351ddb447e6abdede7233a4794d538155.pdf https://github.com/cwfid/dataset
12	University of Bonn Photogrammetry, IGG	Sugar beets dataset for plant classification as well as localization and mapping.	http://www.ipb.uni-bonn.de/data/
13	Flavia leaf dataset	Leaf images of 32 plants.	http://flavia.sourceforge.net/

14	Syngenta Crop Challenge 2017	2,267 of corn hybrids in 2,122 of locations between 2008 and 2016, together with weather and soil conditions	https://www.ideaconnection.com/syngenta-crop-challenge/challenge.php

Reference:

Research papers:

- https://arxiv.org/ftp/arxiv/papers/1807/1807.11809.pdf
- Research Articles
  - https://www.agriculturejournal.org/
  - https://www.nature.com/subjects/agriculture
  - https://juniperpublishers.com/artoaj/