

[← Go Back to Introduction to Computer Vision](#)

:≡ Course Content

Plants Seedling Classification- Problem Statement

Submission type	:	File Upload
Due Date	:	Feb 03, 7:30 AM EST
Total Marks	:	60
Available from	:	Jan 17, 9:30 PM

Description ^

Background and Context

In recent times, the field of agriculture has been in urgent need of modernizing, since the amount of manual work people need to put in to check if plants are growing correctly is still highly extensive. Despite several advances in agricultural technology, people working in the agricultural industry still need to have the ability to sort and recognize different plants and weeds, which takes a lot of time and effort in the long term.

The potential is ripe for this trillion-dollar industry to be greatly impacted by technological innovations that cut down on the requirement for manual labor, and this is where Artificial Intelligence can benefit the workers in this field, *as the time and energy required to identify plant seedlings will be greatly shortened by the use of AI and Deep Learning*. The ability to do so far more efficiently and even more effectively than experienced manual labor could lead to better crop yields, the freeing up of human involvement for higher-order agricultural decision making, and in the long term will result in more sustainable environmental practices in agriculture as well.

Objective

The Aarhus University Signal Processing group, in collaboration with the University of Southern Denmark, has provided the data containing images of unique plants belonging to 12 different species. You being a data scientist, need to build a Convolutional Neural

Network model which would classify the plant seedlings into their respective 12 categories.

Data Description

This dataset contains images of unique plants belonging to 12 different species.

- The data file names are:
 - images.npy
 - Label.csv
- Due to the large volume of data, the images were converted to numpy arrays and stored in images.npy file and the corresponding labels are also put into Labels.csv so that you can work on the data/project seamlessly without having to worry about the high data volume.
- The goal of the project is to create a classifier capable of determining a plant's species from an image.

List of Plant species

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

Guide to solve the project seamlessly

Here are the points which will help you to solve the problem efficiently:

- Read the problem statement carefully
- Download the dataset from the Olympus platform.
- Upload the "images.npy" and "Labels.csv" files to google drive.
- You can set runtime type to "GPU" in Google Colab so that the code will run faster as you will be using CNN to fit your model.

Best Practices for Notebook :

- The notebook should be well-documented, with inline comments explaining the functionality of code and markdown cells containing comments on the observations and insights.
- The notebook should be run from start to finish in a sequential manner before submission.
- It is preferable to remove all warnings and errors before submission.
- The notebook should be submitted as an HTML file (.html) and NOT as a notebook file (.ipynb)

Submission Guidelines :

1. There are two ways to work on this project:

i. Full-code way: The full code way is to write the solution code from scratch and only submit a final Jupyter notebook with all the insights and observations.

ii. Low-code way. The low-code way is to use an existing solution notebook template to build the solution and then submit a business presentation with insights and recommendations.

The primary purpose of providing these two options is to allow learners to opt for the approach that aligns with their individual learning aspirations and outcomes. The below table elaborates on these two options.

Submission type	Who should choose	What is the same across the two	What is different across the two	Final submission file [IMP]	Submission Format

Full - code	Learners who aspire to be in hands-on coding roles in the future focussed on building solution codes from scratch	Perform exploratory data analysis to identify insights and recommendations for the problem	Focus on code writing: 10-20% grading on the quality of the final code submitted	Solution notebook from the full-code template submitted in .html format	.ht ml
Low - code	Learners who aspire to be in managerial roles in the future- focussed on solution review, interpretation, recommendations, and communication with business		Focus on business presentation: 10-20% grading on the quality of the final business presentation submitted	Business presentation in .pdf format with problem definition, insights, and recommendations	.pdf

Please follow the below steps to complete the assessment. Kindly note that if you submit a presentation, ONLY the presentation will be evaluated. Please make sure that all the sections mentioned in the rubric have been covered in your submission.

i. Full-code version

- Download the full-code version of the learner notebook.
- Follow the instructions provided in the notebook to complete the project.
- Clearly write down insights and recommendations for the business problems in the comments.
- Submit only the solution notebook prepared from the learner notebook [format: .html]

ii. Low-code version

- Download the low-code version of the learner notebook.
- Follow the instructions provided in the notebook to complete the project.
- Prepare a business presentation with insights and recommendations for the business problem.
- Submit only the presentation [format: .pdf]

2. Any assignment found copied/plagiarized with other submissions will not be graded and awarded zero marks.
3. Please ensure timely submission as any submission post-deadline will not be accepted for evaluation.
4. Submission will not be evaluated-
 - If it is submitted post-deadline, or,
 - If more than 1 file is submitted.

Best Practices for Full-code submissions

- The final notebook should be well-documented, with inline comments explaining the functionality of code and markdown cells containing comments on the observations and insights.
- The notebook should be run from start to finish in a sequential manner before submission.
- It is important to remove all warnings and errors before submission.
- The notebook should be submitted as an HTML file (.html) and NOT as a notebook file (.ipynb).
- Please refer to the FAQ page for common project-related queries.

Best Practices for Low-code submissions

- The presentation should be made keeping in mind that the audience will be the Data Science lead of a company.
- The key points in the presentation should be the following:
 - Business Overview of the problem and solution approach
 - Key findings and insights which can drive business decisions
 - Business recommendations
 - Focus on explaining the key takeaways in an easy-to-understand manner.
 - The inclusion of the potential benefits of implementing the solution will give you the edge.
- Copying and pasting from the notebook is not a good idea, and it is better to avoid showing codes unless they are the focal point of your presentation.
- The presentation should be submitted as a PDF file (.pdf) and NOT as a .pptx file.

- o Please refer to the FAQ page for common project-related queries.

Happy Learning!

Scoring guide (Rubric) - Plant Seedlings Classification (1) (1)

^

Criteria	Points
Problem Definition and Data Overview - Define the problem statement - Read the dataset properly - Check the shape of the data	6
Perform an Exploratory Data Analysis and get the insights on the images - Plot random images from each of the classes and print their corresponding labels. - Count Plot for each category - Key meaningful observations from EDA	8
Data Pre-processing - Convert the BGR images to RGB images - Resize the images - Plot the images before and after the pre-processing steps - Split the data into train and test - Encode the target variables - Apply the normalization	12
Model building - Build Convolution Neural Network - Evaluate the model on different performance metrics and comment on the performance. For example precision, recall, accuracy. - Plot confusion matrix	8
Model Performance Improvement and Final Model Selection	12

Criteria**Points**

- Build another model using data augmentation to overcome the imbalance problem.
- Evaluate the model on different performance metrics and comment on the performance.
- Plot confusion matrix
- Choose the best model from the ones built with proper reasoning.

Actionable Insights & Recommendations

6

Conclude with the key takeaways for the business

Presentation/Notebook - Overall quality

- Structure and flow
- Crispness
- Visual appeal
- Conclusion and Business Recommendations OR
- Structure and flow
- Well commented code
- Conclusion and Business Recommendations

8

Submit Assignment [Upload file](#) [Previous
Add comments](#)[Next](#)

Proprietary content. ©Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.