



AIML

MODULE PROJECT



1

AIML module projects are designed to have a detailed hands on to integrate theoretical knowledge with actual practical implementations.

2

AIML module projects are designed to enable you as a learner to work on realtime industry scenarios, problems and datasets.

3

AIML module projects are designed to enable you simulating the designed solution using AIML techniques onto python technology platform.

4

AIML module projects are designed to be scored using a predefined rubric based system.

5

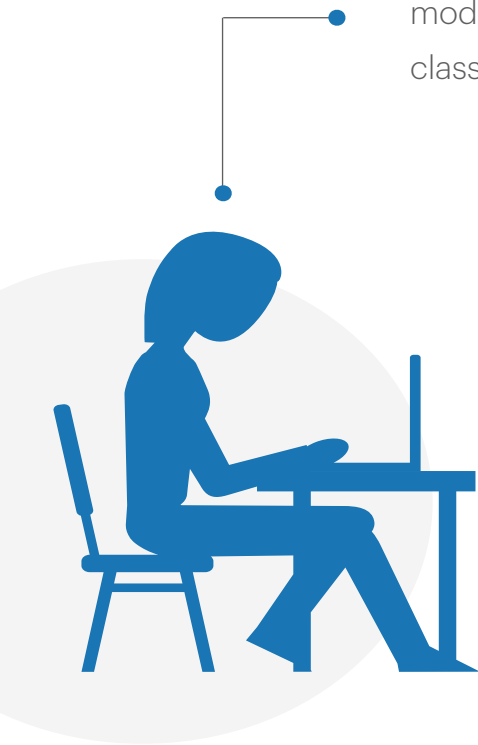
AIML module projects are designed to enhance your learning above and beyond. Hence, it might require you to experiment, research, self learn and implement.

AIML

MODULE
PROJECT

COMPUTER VISION

AIML module project consists of task to train and compare image classifier models using supervised learning classifier, neural network classifier and a CNN classifier.



TOTAL
SCORE

60

PROBLEM STATEMENT

TOTAL
SCORE

60

- **DOMAIN:** Botanical research
- **CONTEXT:** University X is currently undergoing some research involving understanding the characteristics of plant and plant seedlings at various stages of growth. They already have have invested on curating sample images. They require an automation which can create a classifier capable of determining a plant's species from a photo
- **DATA DESCRIPTION:** The dataset comprises of images from 12 plant species. Source: <https://www.kaggle.com/c/plant-seedlings-classification/data>
- **PROJECT OBJECTIVE:** University’s management require an automation which can create a classifier capable of determining a plant's species from a photo

Steps and tasks:

1. Import the data. Analyse the dimensions of the data. Visualise the data.
2. Design, train tune and test the best AIML image classifier model using:
 - Supervised learning algorithms
 - Neural networks
 - CNN
3. Compare the results from the above step along with your detailed observations on the best performing algorithm and why/how it outperforms other algorithms in this case.
4. Pickle the best performing model.
5. Import the the test image [from the “ Prediction” folder] to predict the class. Display the image. Use the best trained image classifier model to predict the class.

LEARNING OUTCOME

Hands-on working knowledge on using ML - supervised learning, NN - Neural networks and CNN - Convolution neural networks for building an image classifier.

In depth working knowledge on working and manipulating images as a data in python

Training, tuning and testing an image classifier.

“Put yourself in the shoes of an actual”

DATA SCIENTIST

THAT'S YOU

Assume that you are working at the company which has received the above problem statement from internal/external client. Finding the best solution for the problem statement will enhance the business/operations for your organisation/project. You are responsible for the complete delivery. Put your best analytical thinking hat to squeeze the raw data into relevant insights and later into an AIML working model.



PLEASE NOTE

Designing a data driven decision product typically traces the following process:

1. **Data and insights:**

Warehouse the relevant data. Clean and validate the data as per the the functional requirements of the problem statement. Capture and validate all possible insights from the data as per the the functional requirements of the problem statement. Please remember there will be numerous ways to achieve this. Sticking to relevance is of utmost importance. Pre-process the data which can be used for relevant AIML model.

2. **AIML training:**

Use the data to train and test a relevant AIML model. Tune the model to achieve the best possible learnings out of the data. This is an iterative process where your knowledge on the above data can help to debug and improvise. Different AIML models react differently and perform depending on quality of the data. Baseline your best performing model and store the learnings for future usage.

3. **AIML end product:**

Design a trigger or user interface for the business to use the designed AIML model for future usage. Maintain, support and keep the model/product updated by continuous improvement/training. These are generally triggered by time, business or change in data.

IMPORTANT POINTERS

Project should be submitted as a single “.html” and “.ipynb” file. Follow the below best practices where your submission should be:

- “.html” and “.ipynb” files should be an exact match.
- Pre-run codes with all outputs intact.
- Error free & machine independent i.e. run on any machine without adding any extra code.
- Well commented for clarity on code designed, assumptions made, approach taken, insights found and results obtained.



Project should be submitted on or before the deadline given by the program office.

Project submission should be an original work from you as a learner. If any percentage of plagiarism found in the submission, the project will not be evaluated and no score will be given.