




2020 US Summer Internship Challenge

Enroll By	Submit By	Result Declaration
 May 15, 2020 1:00 AM	 May 22, 2020 1:00 AM	 May 22, 2020 1:00 AM
<div>\$1</div> <div>1st Place</div>	<div>\$1</div> <div>2nd Place</div>	<div>\$1</div> <div>3rd Place</div>

[DETAILS \(\)](#) [ENROLLEES \(41\) & SUBMISSIONS \(29\) \(\)](#) [RESULTS \(\)](#) [CONVERSATIONS \(1\) \(\)](#) [ATTACHMENTS \(3\) \(\)](#)

Requirements:

Objective

Congratulations,

We have landed on mars and we are trying to colonize it. But the main problem we are facing is which plants we can grow at different geographical regions so that maximum number of plants would survive.

Fortunately we have the data from earth to give us insights.

Please help us in predicting which kind of plant we can grow at different locations depending on the environment.

It's time to save humanity which is at the brink of extinction.

You are our only hope for survival.

Problem Statement

1. Create a report on key insights derived from Exploratory Data Analysis
2. Create a multi-class prediction model to predict the species of plant which will survive in the neighborhood of a given environment.
3. Create key segments for all the plants (train + test) based on the average sunlight

received throughout the day and their distance from waterbody to identify which

segments of plants are getting enough sunlight and water vs which ones are not. This will help in mobilizing resources to track growth of trees appropriately

Key Requirements

Directions -

Hello Challengers,

You will be evaluated based on this Rubric (https://docs.google.com/spreadsheets/d/1_5oK28o4QzvB8Arvo-rD2HsxmCN-I4PPBpiO-mvOY04/edit#gid=2143376223).

Please reach out to AI.Internship@mtxb2b.com (mailto:AI.Internship@mtxb2b.com?subject=Challenge%20Question%20%7C%20%3CFirstName%20LastName%3E) with Subject "Challenge Question | <FirstName LastName>" If you have any questions.

Required sections in the Jupyter Notebook -

1. Exploratory Data Analysis

2. Data Preprocessing**2A. handling outliers (imputation,Removal)****3. Data Engineering****4. Data Preparation for Predictive Modeling****5. Classification Model Predictions (at least 3 different predictive models) with hyperparameter tuning.****6. Comparison of model using performance KPIs, Training & Testing Time****7. Final predictive model recommendation**

Dataset

1. Data size (test and train) : (116203 * 13 and 464809 *13)**2. Target Variable : Plant_Type****3. Data dictionary : Shared separately**

Dataset location

Train Data (<https://drive.google.com/file/d/1reEtaOv3q-qF0PvC2TtDSWV4TnsJCWel/view?usp=sharing>)**Test Data (<https://drive.google.com/file/d/1RidS7W09WZJGvP4IPhFks7uDyHLqsTYZ/view?usp=sharing>)****PRODUCTS**

Challenges

COMPANYLeadership
Blog**RESOURCES**Support
Q&A
Terms of Use
Privacy Policy**NEW CHALLENGES**Development: AI: Document Fraud
Detection
Development: 2020 US Summer Interns
Challenge

(https://www.facebook.com/Maverickio-654095821381664/)



(https://twitter.com/maverick_io)



(https://in.linkedin.com/company/maverick-io)