UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

SUMMER RESEARCH INTERNSHIP REPORT 2022

AT

DALHOUSIE UNIVERSITY-AGRICULTURAL CAMPUS, CANADA

(16th May to 5th August 2022**)**

DALHOUSIE UNIVERSITY

PROJECT TITLE

Nutrient deficiency sensing technology using spectrophotometer

Submitted

DEPARTMENT OF MECHATRONICS ENGINEERING

Name of Student: Harsh Alpeshkumar Senjaliya

Enrolment No.: 201903103510306

Head of Department
(Mechatronics Engineering)
Dr. Chinmay Desai
C.G.Patel Institute of Technology
Uka Tarsadia University

Research Internship Supervisor
Dr. Ahmad Al-Mallahi
McCain Research Chair
AIES Founder and Head
Dalhousie University

Internship external Supervisor/Guide Evaluation form

Purpose: This evaluation is designed primarily to provide feedback on Internship performance and related issues to improve the skills and employability of the student. This form is to be completed and submitted at the end of the Internship by supervisor/guide/mentor under which internship assignment has been executed.

Information about external supervisor/mentor/guide of Intern student			
Name	Ahmad Al-Mallahi	**Contact No**	+1-902.890.2601
Name of Organization	Dalhousie University, Truro	**Email Address**	ahmad.almallahi@dal.ca
Job Title/Designation	Industry Research Chair and Assistant Professor		

Intern Student(s) details					
Enrollment No	**Name of the student**	**Course/Semester**	**Project Title/Internship Name**	**Start date**	**End date**

Intern/Student Evaluation/Assessment Criteria						
Sr. No	**Evaluation criteria**	**Excellent (5)**	**Very Good (4)**	**Satisfactory (3)**	**Needs Improvement (2)**	**Unsatisfactory (1)**
1.	Punctuality and general Behaviors					
2.	Performs in a dependable manner					
3.	Cooperates with co-workers and supervisors					
4.	Shows interest in work					
5.	Learns quickly					
6.	Shows initiative					
7.	Produces high quality work					
8.	Accepts responsibility					
9.	Accepts criticism					

Internship external Supervisor/Guide Evaluation form

10.	Demonstrates organizational skills					
11.	Uses technical knowledge and expertise					
12.	Shows good judgment					
13.	Demonstrates creativity/originality					
14.	Analyzes problems effectively					
15.	Is self-reliant					

Qualitative review/Assessment:

Sr. No	Qualitative Assessment Criteria
a)	**Comment about the approach of Intern/students towards problem solving**
b)	**Suggestions for making internship program more effective and value added**
c)	**Specific observation(s) which will help to improve the quality of intern/student**

Signature of Supervisor/Guide/Mentor	Date

ACKNOWLEDGEMENT

I would like to begin by thanking my internship supervisor honourable Dr. Ahmad Al-Mallahi (Head of AIES) for all of his aid, advice, and directions during the activation phase. I would also want to thank my internship guide, Ms.Reem Abukmeil, for providing me with the assistance and information I needed to submit this internship report as well as the entire AIES laboratory team for being helpful and supportive in every small help I needed and for creating the possibility for me to perform at my best. I am really grateful to Mitacs for providing a research fellowship for the duration of my research internship. I would also like to thank Dr. Manoj Gundalia, the director of the C.G.Patel Institute of Technology and Dr.Chinmay Desai the head of the Mechanical and Mechatronics Engineering Department for allowing me to prolong my internship. Finally, I would want to thank Dr. Ghanshyam Barman and Assistant Professor Mrs.Nita Patil for their help in preparing for this internship, as well as everyone else involved.

TABLE OF CONTENT

I. **Overview of University:**

Dalhousie University-Agricultural Campus, Truro, Nova Scotia, Canada

The University Campus, located just outside Truro in Bible Hill, Nova Scotia, is home to a thriving farm, almost 1,000 acres of research fields, gardens, and greenhouses, and has a long heritage of industry-leading teaching and research dating back to 1905. Every year, Dalhousie welcomes over 20,000 students to their campuses and celebrates the accomplishments of over 120,000 alumni.

The university was rated in the top 300 in the QS World University Rankings in 2021, and 12th in Canada. In addition, it was rated 251-300th in the Times Higher Education World University Rankings, and it was in the top 15 in Canada. The Shanghai Ranking's Academic Ranking of World Universities placed Dalhousie in the top 300 globally and in the top 12 in Canada.

Agriculture is an exciting and hard business in today's world where you may make a meaningful difference. This university's curricula are unusual in that they blend the basic disciplines of agricultural sciences with business management and cutting-edge technology.

About Research Group:

Applied Intelligent Engineering Systems (AIES) Research Group

The purpose of the AIES research group is to design engineering systems that support sustainable agricultural operations and contribute to regenerative crop production, with an emphasis on field application-driven research that meets the demands of producers and processors. The Group is headed by Dr.Ahmad Al-Mallahi (McCains food Research Lead). McCain Foods Limited and Potatoes New Brunswick (PNB) are funding the research, and by focus on difficulties throughout the potato production chain, such as precision planting, spot spraying, autonomous scouting, post-harvest monitoring, and storage management. AIES’s research initiatives are currently supported by NSERC's CRD program, Mitacs' Accelerate program, and AAFC's Canadian Agricultural Partnership.

II. Project Title

Nutrient Deficiency Sensing Technology using Spectrophotometry

Project Goals

To develop nutrient estimation sensing technique for potato plants based on spectroscopy

Intern's Activities

The development includes data collection from the field and indoor. Data collection includes manually sampling potato petioles and leaves, learning how to dry and transport them, and tabulating data received from the laboratory. The activities include learning how to use portable spectrophotometer and arrange raw data. Also the activities include writing software codes to manipulate the data on R or Python languages and learning how to test statistical models for validation.

Deliverables

I. Monitoring the growth of plants indoor.
II. Building datasets of nutrient concentrations and spectral reflectance.
III. Testing and validating developed nutrient estimation models.

Project Objective

Variable rate technology aims to precisely apply inputs to the farm to optimise the usage of input, reduce cost, and save environment. It requires information about the land status to be able to respond to variations. Although the concept has been discussed in literature, the expansion of the technology has not been wide yet. One of the problems is lack of suitable sensing technology for many inputs. In this project, the objective is to develop a sensor to detect symptoms of nutrient deficiency in plants. There are certain symptoms that appear on the leaves of plants when they suffer from nutrient deficiency such as changing colour or shape. Early detection of deficiency will allow the farm manage to respond so that losses in yield quantity and quality is avoided. We plan to scan the field by cameras searching for symptoms and information about symptoms will be synchronised with GPS information to create nutrient deficiency maps. Besides developing a novel sensor, the sensing system should include methods to physically scan the fields in a comprehensive, fast, and accurate manner, as well as considering the external effects on the sensor when working outdoors.

III. **Introduction**

The goal of this Internship was to create a sensing technology to assess nutrients in potato plants that can enhance current scouting procedures by decreasing the period between data collection and action and lowering data collecting expenses. Non-destructive ways for providing efficient information on plant nutritional content utilising leaf/canopy reflectance have been introduced to the market. Ground-based and remote sensors are frequently employed to identify stressed plants by acquiring electromagnetic wave reflectance information from the canopy. However, both sensors have disadvantages in terms of canopy reflectance, including as air and soil interferences. Reflectance at the leaf level has received a lot of attention as a way to get rid of such sounds. Those ground-based sensors at the leaf level have been installed. Those ground-based sensors at the leaf level have been evaluated for their ability to assess nutrients using the chemical composition of the leaf as the ground truth. Petioles, rather than leaf chemical testing, are the major organ for tissue testing in potato plants. Collecting samples, preparing the data that had been obtained, and performing non-linear regression were the overall objectives for this internship.

IV. **Data Collection Methodology**

IV**a**.Indoor Data Collection Methodology

Managed autonomous indoor planting system at Banting Annex building situated on Dalhousie University, Truro campus, which control of temperature and artificial light. We managed deployed sensors in the room to give a good indication of the status of humidity, temperature and light intensity using SHT sensor and Apogee SQ515 sensor. The following sensors were interfaced with Microcontroller (Arduino UNO) which was further interfaced with Raspberry -Pi where sensed inputs triggers the connected Actuators (Exhaust Fan, Heater and Artificial Lights) to regulate indoor environment.

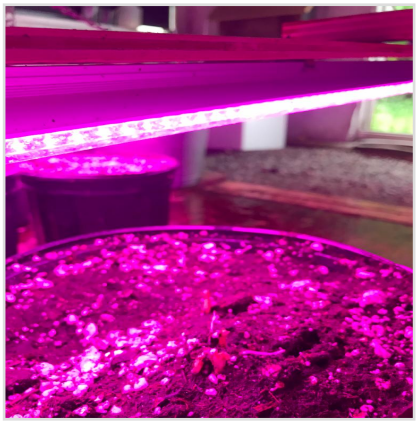


Fig1: Potato Plant Pods Under Artificial Light at Indoor planting facility

Fig 2: Hardware interface at Indoor Planting Facility

Plantation of the Russet Burbank type of potato was done with help of every AIES members.The most popular potato in North America is the Russet Burbank cultivar, which has dark brown skin and few eyes. We picked seed potatoes from pre-sprouted potatoes. As seed potatoes should be approximately 100 grams, we sliced huge tubers into sections and measured them on a digital weight lab scale. We sowed seed potato segments sliced side down (eyes up) in tubs filled with previously watered pro-mix mycorrhizae growing medium.Preparing sample package to perform the chemical

Fig 3: Indoor Potato Plantation Preparation

analysis on the indoor grown potato. The instructed collection of petioles and leafs was fourth from the top of the plant as suggested by potato experts at Florenceville-Bristol. Each sample package contains 40 leaves and 20 petioles samples for chemical and spectral analysis.

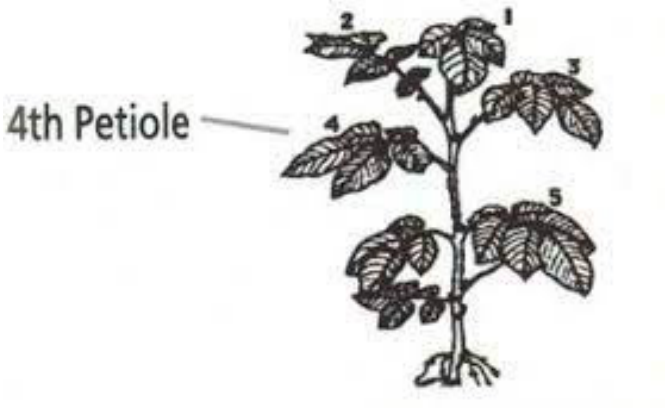


Fig 4.: Image indicating 4th petiole of potato

IV**b**.Field Data Collection Methodology

Data acquired on open potato fields differed significantly from data collected indoor. To perform this task we visited several potato fields near Florenceville-Bristol and Grand falls to collect spectral data and data for chemical analysis.

The Spectral analysis of potato plant nutrient were carried out with the help of spectrophotometer device named "**ASD FieldSpec 4**" as showed in Fig.5 which provides spectral reading from 350nm-2500nm wavelength. Higher resolution hyperspectral sensors yield greater precision for remote sensing classification applications, producing more information from every pixel generated in an image than ever before. Using this mentioned spectrophotometer device we collected spectral data on the potato fields at fixed timings between 9:00 AM to 12:00 Noon. As in the mentioned time position of sun provides best spectral results.

Fig 5: ASD FieldSpec

Two strategies are used in the data gathering operation. I) Data on dry leaves; II) Data on fresh leaves.

We developed packages called field unique identification numbers while gathering data on fields as illustrated in Fig.6. The sample gathered from the specific field is placed in a unique numbered envelope and later dried. The purpose of this activity is to remove water from the leaves, which aids in the reduction of noise in the spectral data. Each envelop contains 40 leaves and 20 petioles which makes one data point.

Fig. 5: Sample Collection Method

V. **Performing Data Preprocessing on the collected dataset**

V**a**.Principle Principal component Analysis [PCA]

Large datasets are becoming more prevalent; however, they are frequently challenging to comprehend. Principal component analysis (PCA) is a technique for lowering the dimensionality of such datasets, boosting interpretability while minimising information loss. It accomplishes this by generating new uncorrelated variables that successively optimise variance. Implemented Principal Component Analysis (PCA) on the dataset collected in 2020 (Boron).

V**b**.Spearmen Correlation Analysis [SCA]

The Spearman correlation coefficient/Analysis (SCC/SCA) is a nonparametric variation of the Pearson correlation coefficient that calculates the degree of relationship between two variables based on their ranks. Often denoted by the Greek letter ρ (rho) and r_s.

The Pearson Product Moment Correlation examines the relationship between two continuous variables. A linear relationship is one in which two variables change in the same direction at the same pace.

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

where,
ρ = Spearman Correlation coefficient
$\mathbf{d_i}$ = the difference in the ranks given to the two variables values for each item of the data,
n = total number of observations

The variables may have a positive relationship, which means that when one variable's value increases, so do the values of the other variable. The relationship can also be negative, which means that when the values of one variable rise, the values of the others fall. Finally, the correlation could be neutral, implying that the factors are unrelated.

Correlation quantifies this link, typically as a value ranging from -1 to 1 for totally negatively correlated and perfectly positively correlated data as shown in Fig.6. The resulting correlation is known as the "correlation coefficient." This correlation coefficient can then be used to interpret the measures.

Fig. 6: SC Analysis on spectral bandwidths chosen by Lasso regression

VI. **Non-linear Regression on Dataset collected**

A dataset was gathered in 2020, and linear regression analysis (Lasso) was performed. Lasso regression is a linear regression technique that employs shrinkage. In the validation findings, the Lasso model r^2 value for Boron was 0.62. We tested the results of r^2 using non-linear regression methods such as Logarithmic regression, Exponential regression, Logistic regression, and Random Forest regression on the same dataset.

In the mathematics and social sciences, as well as in engineering, nonlinear regression analysis is a widely used approach. The term "nonlinear regression models" refers to models with nonlinear parameters. Nonlinear regression analysis is primarily concerned with response prediction, statistical inferences of parameter estimations, and the nonlinear model's quality of fit.

The dataset acquired in 2020 (Dried leaves; Boron) was used for the non-linear regression analysis. We structured the test into three sections:

1. In the model, Lasso chosen data was implemented.
2. Implemented PCA's data.
3. Implemented a merged collection comprising data collected both indoors and outside (Dried leaves).

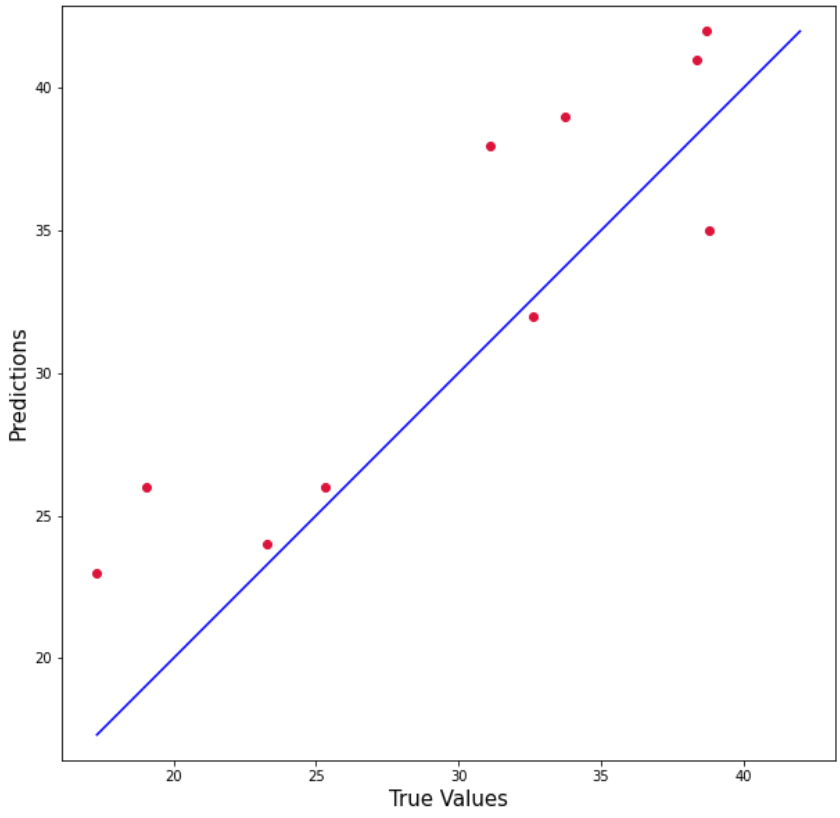
VI**a**.Logarithmic Regression

Logarithmic regressions are commonly used to model events that expand rapidly at beginning, slow down with time, yet continue to grow unboundedly. Here we have implemented Logarithmic Regression analysis on the 2020 dataset (Dried (Boron)). The produced results are shown below.

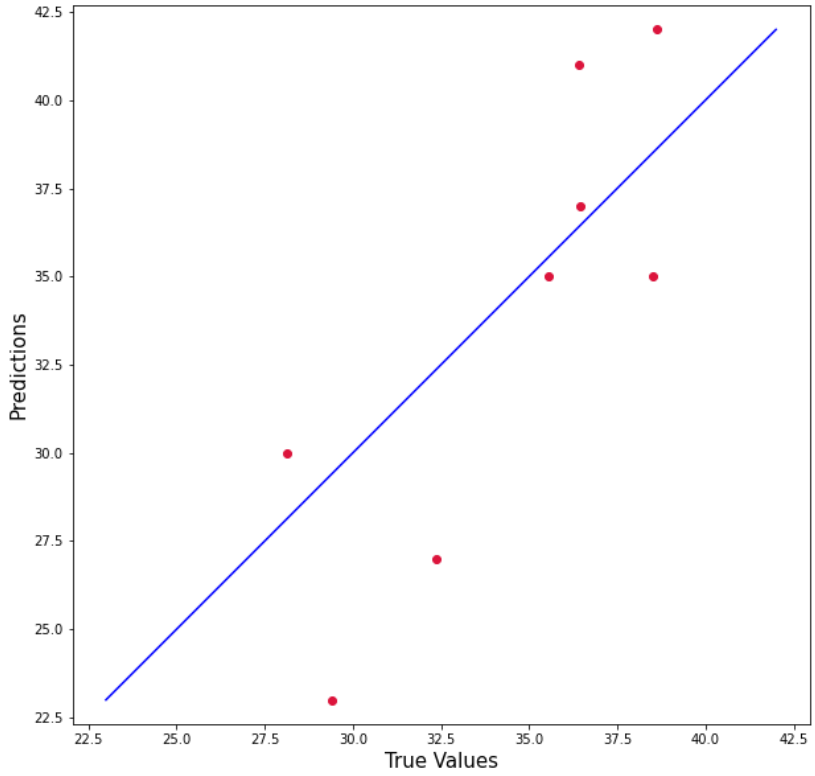
Equation: $y = a + b*\ln(x)$

Note: The model was trained and tested at an 80:20 ratio.

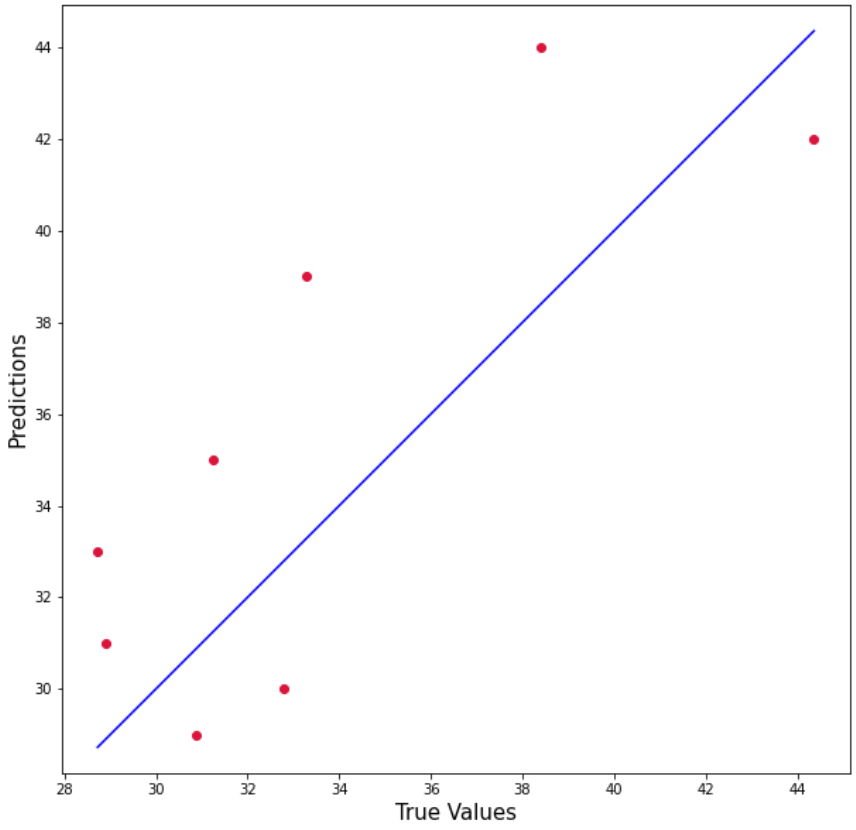
Test-1: Lasso chosen data was implemented.



Test-2: Implemented PCA's data



Test-3: Implemented a merged collection comprising data collected both indoors and outside (Dried leaves).



TEST NO.	R^2 VALUES
Test-1	0.46
Test-2	0.57
Test-3	0.4525

Table 1: R^2 Value (Logarithmic Regression)

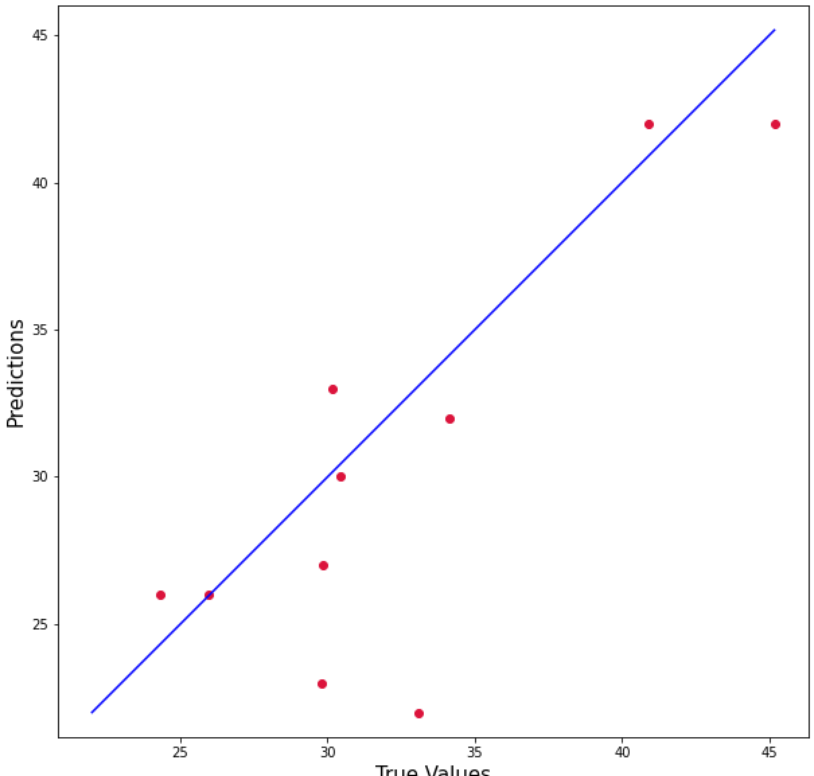
VI**b.**Exponential Regression

Exponential regressions are typically used on phenomena where the growth begins slowly and then increases very rapidly as time increases. Here we have implemented Exponential Regression analysis on the 2020 dataset (Dried (Boron)). The produced results are shown below.

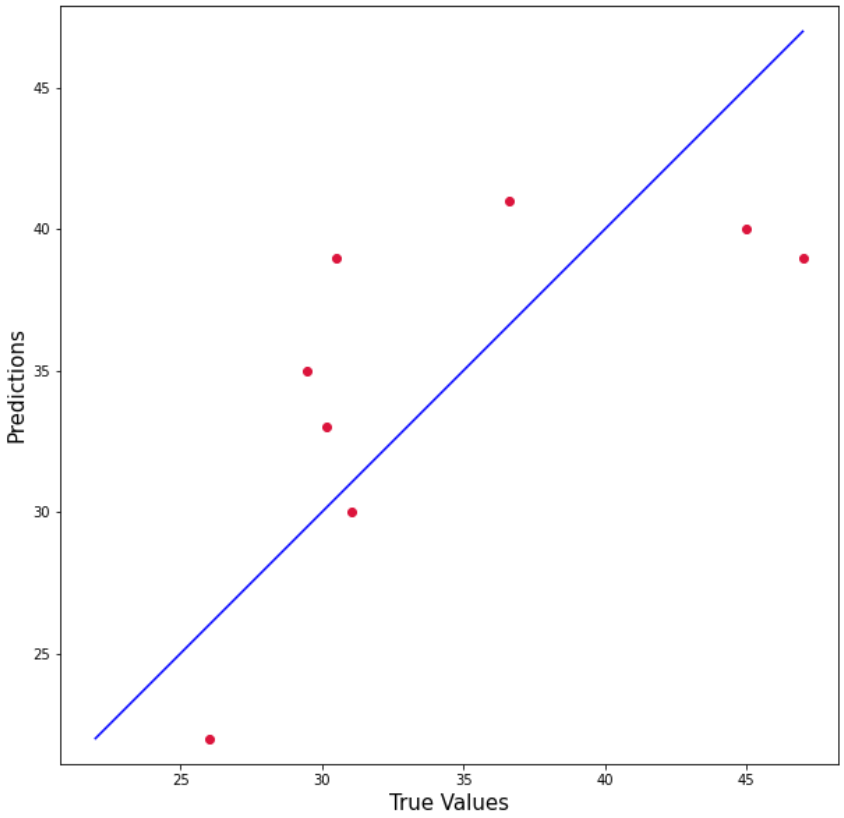
Equation: $y=ab^x$

Note: The model was trained and tested at an 80:20 ratio.

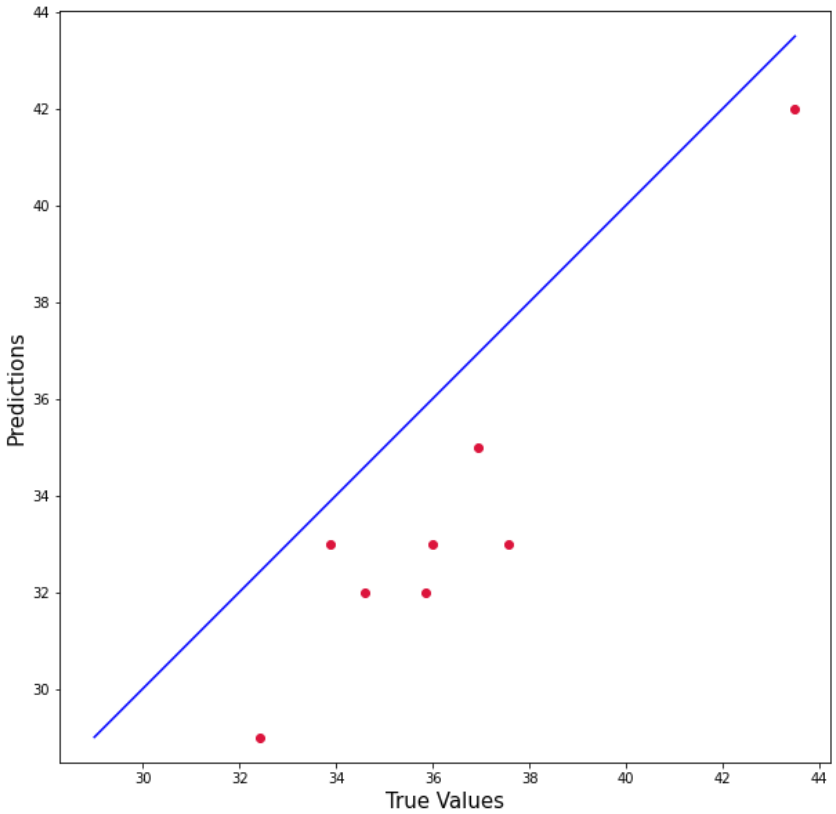
Test-1: Lasso chosen data was implemented.



Test-2: Implemented PCA's data.



Test-3: Implemented a merged collection comprising data collected both indoors and outside (Dried leaves).



TEST NO.	R^2 VALUES
Test-1	0.41
Test-2	0.38
Test-3	0.2308

Table 2: R^2 Value (Exponential Regression)

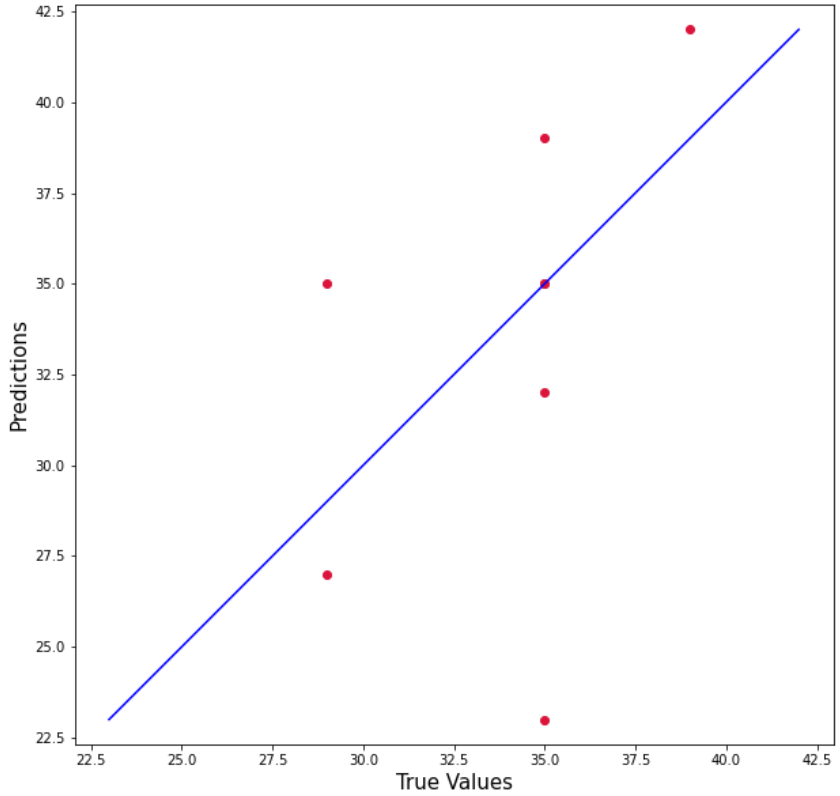
VI**c.**Logisitic Regression

Logistic regression models the probabilities for classification problems with two possible outcomes. It's an extension of the linear regression model for classification problems. Here we have implemented Logistic Regression analysis on the 2020 dataset (Dried (Boron)). The produced results are shown below.

Equation: logistic(η)=1/1+exp(−η)

Note: The model was trained and tested at an 80:20 ratio.

Test-2: Implemented PCA's data.



TEST NO.	R^2 VALUES
Test-2	0.1742

Table 3: R^2 Values (Logistic Regression)

VI**d.**Ranadom Forest Regression

Random Forest Regression is a supervised learning technique that does regression using the ensemble learning method. The ensemble learning approach combines predictions from numerous machine learning algorithms to provide a more accurate forecast than a single model. Here we have implemented Random Forest Regression analysis on the 2020 dataset (Dried (Boron)). The produced results are shown below.

Note: The model was trained and tested at an 80:20 ratio.

Test-1: Lasso chosen data was implemented.

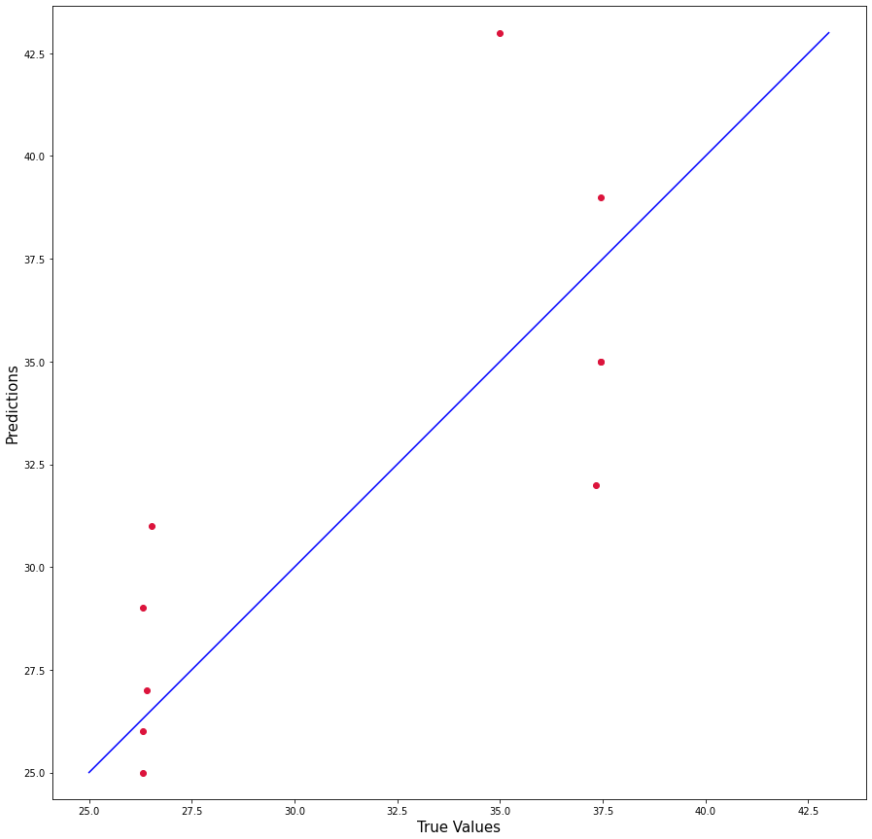
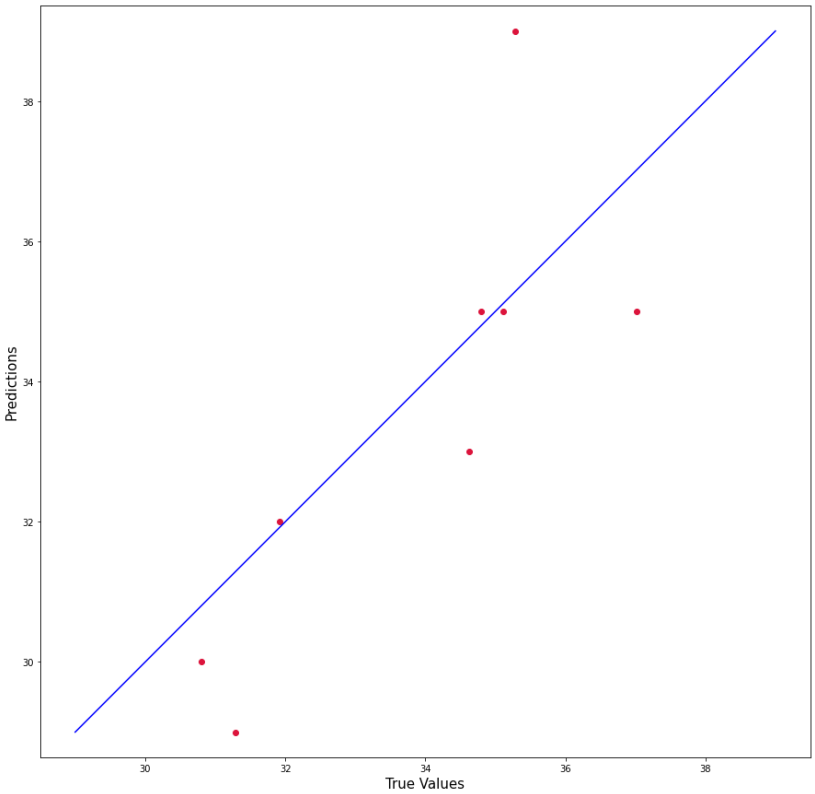
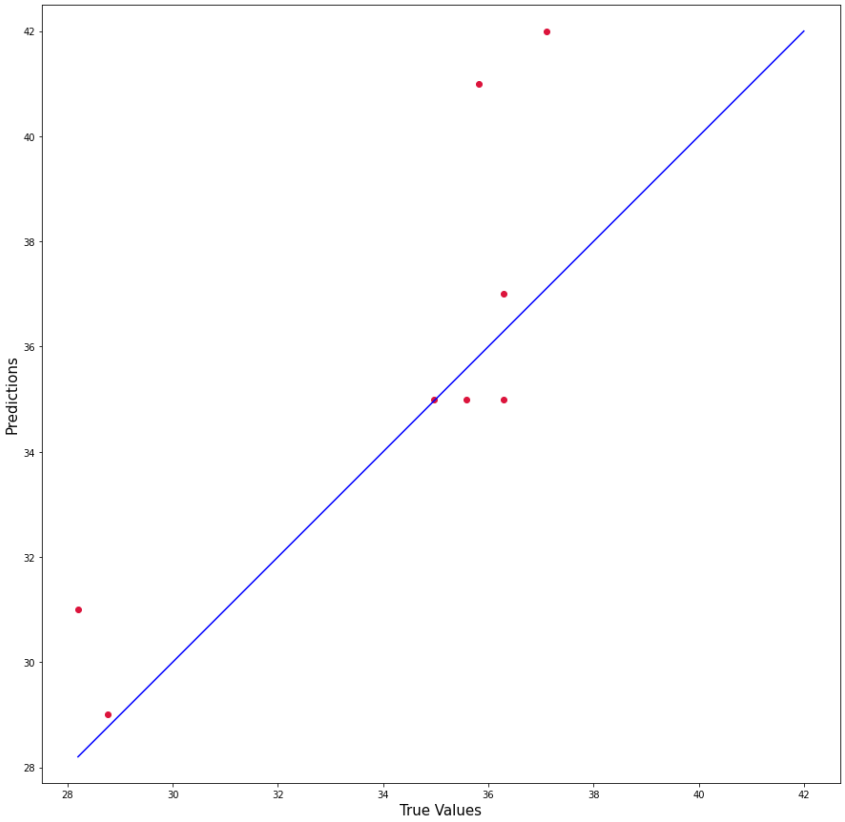


Table 4: R^2 Values (Random Forest Regression)

Test-2: Implemented PCA's data.



Test-3: Implemented a merged collection comprising data collected both indoors and outside (Dried leaves).



TEST NO.	R^2 VALUES
Test-1	0.5566
Test-2	**0.6324**
Test-3	0.4080

Miscellaneous Tasks

Learned to Develop communication between Raspberry Pi and CANBus module from Ph.D. student Mozammel Motalab and Master Arshvir Singh. This communication system is used to interface and establish communication between virtual terminal system to the static boom sprayer controlling system. Along with this also got opportunity to learn integrating GPS with prescription map and operate the nozzles to spray accordingly on the static sprayer as illustrated in Fig. 7.

Fig. 7:Operating nozzles through virtual terminal and integrating GPS with prescription map.

Conclusion

The above mentioned tasks provided depth knowledge of to perform nutrient analysis of potato plant which contains, data collection, data preprocessing, raw data analysis, data handling and data analysis using machine learning techniques. During the tenure, I gained knowledge about spectrophotometer device and collect spectral data. Also got depth knowledge of arranging apparatus, handling the instruments and conducting fair research. This twelve week of research opportunity provided enormous exposure to the scientific universe and community. My Key learning from this research internship is, research is all of trying rather than a good result. I am thankful to Dr. Ahmad Al-Mallahi and Ms.Reem Abukmeil to provide best equipment to conduct research and nurturing during the internship period.