Increasing yield of Crop Production in Greenhouses using AI

Capstone Project Team – Group 5

Yellai Aditya Venkat Murty – 0801899

Odunayo Agagu – 0810500

Khushi Patel – 0790434

Suman Sunny Rathod – 0815907

Naveen Kumar Reddy Sama – 0804315

Agenda

Problem Statement

Dataset Information

Methodology

Dashboards

Key Insights

Data collection

Greenhouse Visit

Machine Learning Models

Use Case

Challenges

Problem Statement

Issue

The agriculture sector is faced with challenge of optimizing yield and managing crop disease amidst increasing population and dwindling resources. Greenhouses – major form of crop growing environment in Canada will be the focal point of consideration.

Concerns

Crop Damage, Resource inefficiency, environmental damage and economic loss resulting from surplus/waste.

Goal

Crop Analyze production, current practices and major crops grown in greenhouses across Canada.

Optimize issue faced with respect to planting, managing diseases in crop to optimize yield Help adopt best practices suitable for countering disease infestation with efficiency and speed.

Dataset Information

Source

https://www150.statcan.gc.ca/
https://universe.roboflow.com/search

Relational Dataset

Greenhouses: area covered, major crops grown, yield, import & export

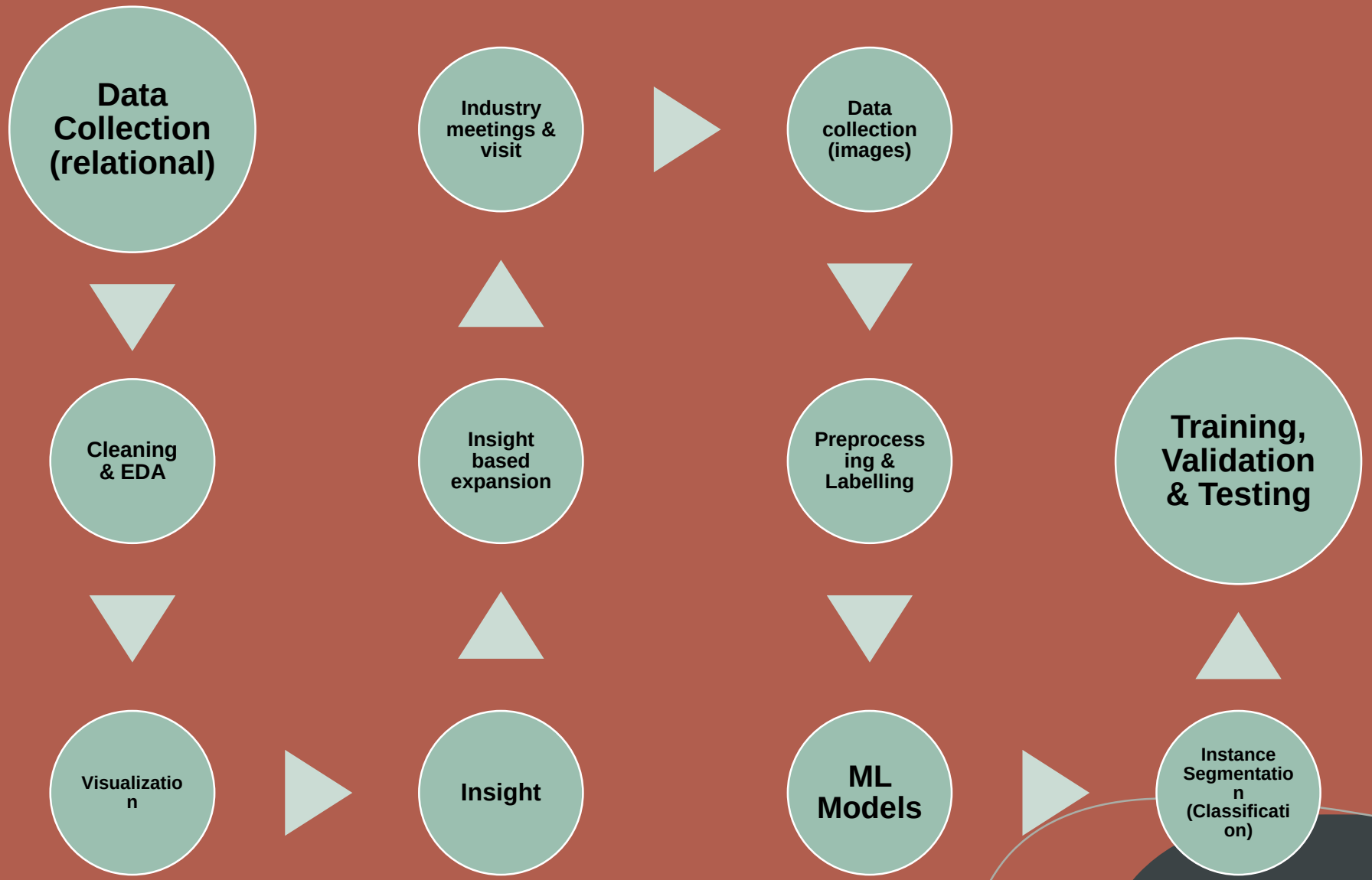
Image Dataset

Five Identified major crop disease images (leaves)

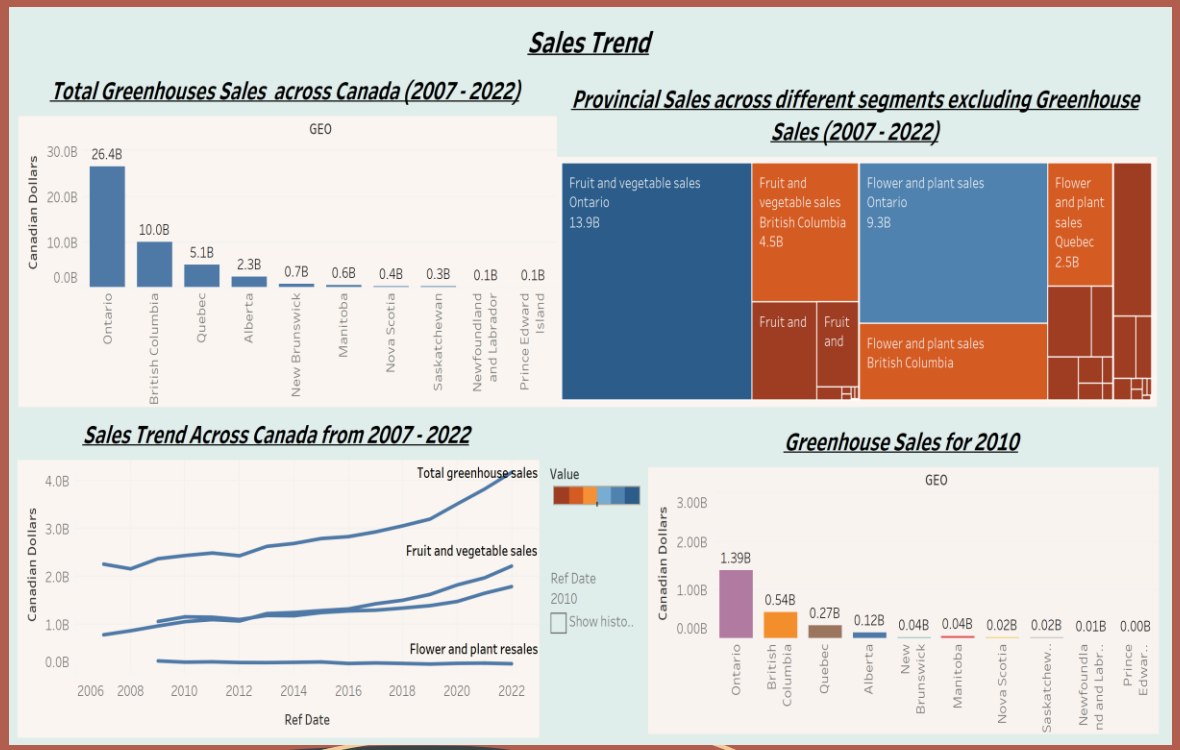
Database

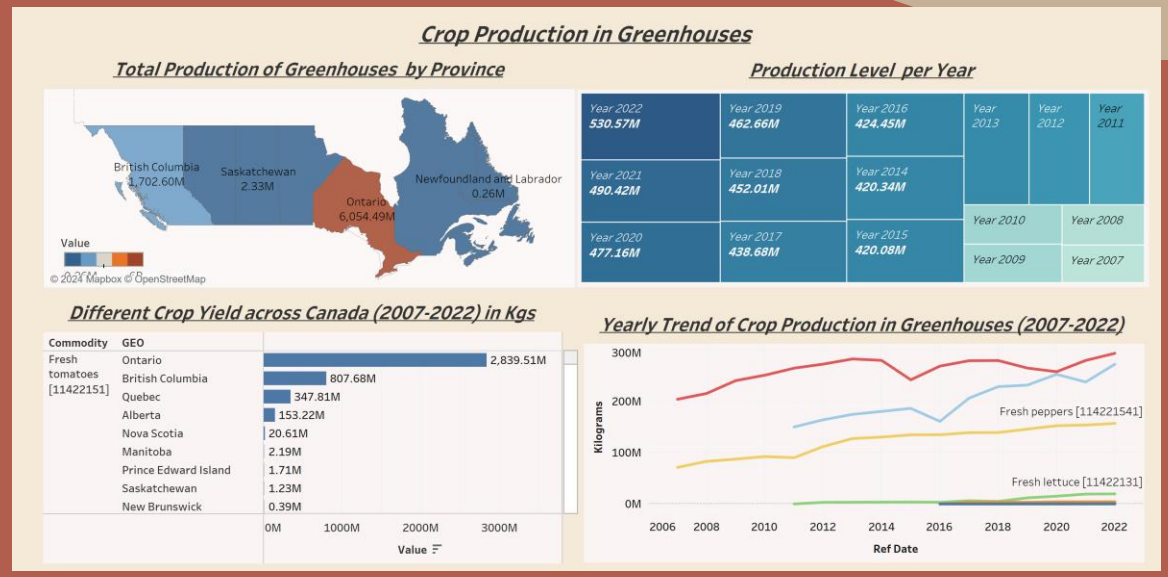
Phase 1: Excel fed into tableau
Phase2: JPG images managed via Roboflow

Methodology



Dashboards


Sales Trend
Total Greenhouses Sales across Canada (2007 - 2022)
GEO
Canadian Dollars
30.0B
20.0B
10.0B
0.0B
26.4B
10.0B
5.1B
2.3B
0.7B
0.6B
0.4B
0.3B
0.1B
0.1B
Ontario
British Columbia
Quebec
Alberta
New Brunswick
Manitoba
Nova Scotia
Saskatchewan
Newfoundland and Labrador
Prince Edward Island
Provincial Sales across different segments excluding Greenhouse Sales (2007 - 2022)
Fruit and vegetable sales Ontario 13.9B
Fruit and vegetable sales British Columbia 4.5B
Flower and plant sales Ontario 9.3B
Flower and plant sales Quebec 2.5B
Flower and plant sales British Columbia
Fruit and
Fruit and
Sales Trend Across Canada from 2007 - 2022
Total greenhouse sales
Fruit and vegetable sales
Flower and plant resales
Canadian Dollars
4.0B
3.0B
2.0B
1.0B
0.0B
2006
2008
2010
2012
2014
2016
2018
2020
2022
Ref Date
Value
Ref Date
2010
Show histo..
Greenhouse Sales for 2010
GEO
Canadian Dollars
3.00B
2.00B
1.00B
0.00B
1.39B
0.54B
0.27B
0.12B
0.04B
0.04B
0.02B
0.02B
0.01B
0.00B
Ontario
British Columbia
Quebec
Alberta
New Brunswick
Manitoba
Nova Scotia
Saskatchew...
Newfoundla nd and Labr..
Prince Edwar..


Crop Production in Greenhouses
Total Production of Greenhouses by Province
British Columbia 1,702.60M
Saskatchewan 2.33M
Ontario 6,054.49M
Newfoundland and Labrador 0.26M
Value
© 2024 Mapbox © OpenStreetMap
Production Level per Year
Year 2022 530.57M
Year 2019 462.66M
Year 2016 424.45M
Year 2013
Year 2012
Year 2011
Year 2021 490.42M
Year 2018 452.01M
Year 2014 420.34M
Year 2010
Year 2008
Year 2020 477.16M
Year 2017 438.68M
Year 2015 420.08M
Year 2009
Year 2007
Different Crop Yield across Canada (2007-2022) in Kgs
Commodity
GEO
Fresh tomatoes [11422151]
Ontario 2,839.51M
British Columbia 807.68M
Quebec 347.81M
Alberta 153.22M
Nova Scotia 20.61M
Manitoba 2.19M
Prince Edward Island 1.71M
Saskatchewan 1.23M
New Brunswick 0.39M
0M
1000M
2000M
3000M
Value
Yearly Trend of Crop Production in Greenhouses (2007-2022)
Kilograms
300M
200M
100M
0M
Fresh peppers [114221541]
Fresh lettuce [11422131]
2006
2008
2010
2012
2014
2016
2018
2020
2022
Ref Date


Land Area of Greenhouses across Canada (2007 - 2022)
Area of Greenhouses per Province (Sq.m)
GEO
Ontario 15.95M
British Columbia 4.99M
Quebec 2.65M
Alberta 1.23M
Manitoba 0.31M
Saskatchewan 0.13M
Nova Scotia 0.18M
New Brunswick 0.15M
Newfoundland and Labra.. 0.07M
Prince Edward Island 0.05M
Types of Greenhouses and Land Area
Area
GEO
Poly-film
Ontario 9.3M
Quebec 2.1M
British Columbia 0.9M
Alberta 0.8M
Manitoba 0.3M
New Brunswick 0.1M
Saskatchewan 0.1M
Nova Scotia 0.1M
0.0M
2.0M
4.0M
6.0M
8.0M
10.0M
Square Metres
Total Greenhouses in Operation - 2019
GEO
Count
1.5K
1K
0.5K
0K
1,110
830
630
275
130
139
130
80
62
24
Ontario
Quebec
British Columbia
Alberta
Saskatchewan
Manitoba
Nova Scotia
New Brunswick
Newfoundland and Labrador
Prince Edward Island
Value
0.03M
19.62M
Ref Date
2019
Show history

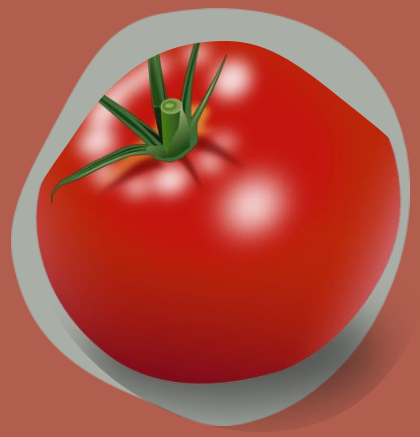
Key Insights

The three (3) provinces that account for the highest greenhouses (in terms of number, total area, production and sales) in Canada are Ontario, British Columbia and Quebec.

The major commodities produced in greenhouses are tomatoes, cucumbers and peppers. Followed by lettuce, strawberries and eggplants.

The greenhouses are either made of poly-film, glass or rigid plastic. Poly-film made greenhouse cover the large land area

Image Data Collection - Crops



Tomato

Cucumber

Strawberry

Lettuce

Bell Pepper

Data Collection – 5 crops, 2 major diseases each. (Collect Images from Roboflow/google)

Classify – 1. Healthy / Unhealthy (Binary) 2. Disease classification (multi)

Preprocessing – sort in folder, remove duplicates, rename and resize

Labelling – Manual annotation, highlight infected segments in images for training

Augmentation – Use robolow to complete augmentation of labelled dataset

Greenhouse Visit Insights

After a keen discussion with the greenhouses, this modelling can help reduce the disease detection time & improve the quality of produce.

Image classification can help set the disease type into different stages helping the greenhouses to act differently by sending them to the lab for examination and understand which has to be treated in what way.

Overall, will boost efficiency, reduce labor and time consumed, and can act faster to reduce damage due to disease leading to crop yield optimization.

Machine Learning Models

CNN

Ability to analyze spatial patterns in image data, important for disease detection

VGG16

Deep architecture and proven effectiveness in image classification tasks, enhanced accuracy.

Yolov8s

Most suitable for efficient classification involving segmentation, identifying anomalies. Enormous Computational capacity and augmentation

Roboflow: Yolov8sInstance Segmentation

Beyond simple image classification – not only identifies objects but also precisely labels individual instance of object within image.

Distinguishes between different instances of same class. For example: not only class dogs in image also segment different dog breeds.

Detailed object recognition, precise identification using enhanced visual insights

CNN & VGG16

CNN (Binary Classification)

Accuracy: 0.97

VGG16 (Multi Classification)

Accuracy: 0.33

VGG19 (Multi Classification)

Accuracy: 0.47

Model: Yolov8s

mAP	Precision	Recall
42.3%	49.4%	39.9%

Model Type: Roboflow 3.0 Instance Segmentation (Fast)
Checkpoint:crop-disease-identification/8
Version: 9
Data Split Ration: 70:20:10

Use cases

Early Disease Detection

Early detection due to automation enables timely and targeted action

Resource Management

Better resource allocation including water, nutrients, pesticides and human resource

Yield Optimization

Approximate more efficiently on crop yields and plan accordingly

Quality Control

Less disease and anomalies ensures better quality

Challenges

Data Collection

Required variables in a dataset crucial for crop yield prediction and/or optimization (soil nutrients, pesticides, and insecticides) were confidential.

Multiple attempts from multiple channels did not yield result for what was required.

Images collected had quality issues that had to be tackled efficiently.

Computational Limitations

Image dataset after augmentation resulted in huge dataset requiring proper pipeline and memory.

Code prepared to be run on python/google colab could not run all required epochs due to computational limit.

Less epochs created poor accuracy and generalization.

Industry Knowledge

Considerable amount of time was spent on research and meeting industry experts including site visit to navigate through best practices and requirements to navigate through the project.

Thank You

References

1. https://www.sciencedirect.com/science/article/pii/S0168169920302301
2. https://www.nature.com/articles/s41598-023-42843-2
3. https://www.sciencedirect.com/science/article/pii/S0168169920302301#:~:text=Several%20machine%20learning%20algorithms%20have,in%20crop%20yield%20prediction%20studies
4. https://docs.roboflow.com/datasets/image-augmentation
5. https://www.statcan.gc.ca/en/data-science/network/greenhouse-detection
6. https://docs.ultralytics.com/#yolo-licenses-how-is-ultralytics-yolo-licensed
7. https://yolov8.com/