**HARVESTIFY**

**Project Report**

**Version 1.0**

Project Work Phase – II (EAI852)

**BACHELOR OF TECHNOLOGY (CSE)**

**SPECIALIZATION-**

**ARTIFICIAL INTELLIGENCE**

**MACHINE LEARNING**

**AND DEEP LEARNING**

Project Guide: Submitted By:

Mr. Ashish Bishnoi Atishay Jain (TCA1959018)

Assistant Professor Charu Saxena (TCA1959050)

CCSIT, TMU

JAN, 2023

**FACULTY OF ENGINEERING & COMPUTING SCIENCES**

**TEERTHANKER MAHAVEER UNIVERSITY, MORADABAD**

**DECLARATION**

We hereby declare that this Project Report titled **Harvestify** submitted by us and approved by our project guide, Faculty of Engineering & Computing Sciences. Teerthanker Mahaveer University, Moradabad, is a bonafide work undertaken by us and it is not submitted to any other University or Institution for the award of any degree diploma / certificate or published any time before.

|  |  |  |
| --- | --- | --- |
| **Project ID :** |  | |
| **Student Name:** | Atishay Jain |  |
| **Student Name:** | Charu Saxena |  |
|  |  |  |
| **Project Guide:** | Mr. Ashish Bishnoi |  |

Table of Contents

[1 Project Title 4](#_Toc135493245)

[2 Problem Statement 4](#_Toc135493246)

[3 Project Description 5](#_Toc135493247)

[3.1 Scope of the Work 6](#_Toc135493248)

[3.2 Project Modules 7](#_Toc135493249)

[3.3 Context Diagram (High Level) 13](#_Toc135493250)

[4 Implementation Methodology 14](#_Toc135493251)

[5 Technologies to be used 16](#_Toc135493252)

[5.1 Software Platform 16](#_Toc135493253)

[5.2 Hardware Platform 17](#_Toc135493254)

[5.3 Tools, if any 17](#_Toc135493255)

[6 Advantages of this Project 18](#_Toc135493256)

[7 Assumptions, if any 18](#_Toc135493257)

[8 Future Scope and further enhancement of the Project 19](#_Toc135493258)

[9 Project Repository Location 20](#_Toc135493259)

[10 Definitions, Acronyms, and Abbreviations 21](#_Toc135493260)

[11 Conclusion 21](#_Toc135493261)

[12 References 22](#_Toc135493262)

**Appendix**

**A: Data Flow Diagram (DFD)**

**B: Use Case Diagram (UCD)**

**C: Screen Shots**

# Project Title

Harvestify

# Problem Statement

Currently, agriculture is facing a hideous problem. Where in spite of all right knowledge consumed, an agriculture sector is facing a huge loss. Why? To put simply the crops aren’t supervised properly. classification analysis can help you with finding the right area for your crop, further resulting to control damage and more revenue generation. To be able to successfully yield crops, foremost and major key role is proper irrigation functionality. Machine learning algorithm can help with better irrigation resulting in following ways – maintain a desired soil water range in the root zone that is optimal for plant growth, low labor input for irrigation process management.

Farming is about risk calculation but what if the risk can be calculated and cured beforehand. Anomaly analysis can help you with identifying the weakness and strength of the soil, resulting in more revenue generation and saving ample amount of time.

Climate is now a data problem,” says Claire Monteleone. Earlier, Improper weather predictions lead to many crops lost- resulting in loss of money and time invested. But technology has evolved over years leading businesses to higher stable growth. Regression analysis will help you with better production forecasting using weather**.** conditions.Farming is one of the major sectors that influences a country’s economic growth.

In country like India, majority of the population is dependent on agriculture for their livelihood. Many new technologies, such as Machine Learning and deep Learning, are being implemented into agriculture so that it is easier for farmers to grow and maximize the yield.In this project, I present a website in which the following applications are implemented; Crop recommendation, Fertilizer recommendation and plant disease prediction, respectively.

* In the crop recommendation application, the user can provide the soil data from their sideand the application will predict which crop should the user grow.
* For the Fertilizer recommendation application, the user can input the soil data and the type of the crop they are growing, and the application will predict what the soil lacks or has excess of and will recommend improvements.
* For the last application, that is the plant disease prediction application, the user can input an image of a diseased plant leaf, and the application will predict what disease it is and will also give a little background about the disease and suggestion to cure it.

The works done till now only concentrated on crop prediction using different soil properties and Data Mining Techniques. Fertilizer Recommendation is not taken into consideration. So, it is necessary to develop crop yield prediction and fertilizer recommendation system which predicts crop yield based on soil nutrients crop yield data and recommend fertilizer for selected crop based on different datasets like fertilizer data, location data and crop yield data.

The Harvestify Machine learning project aim to develop and implement intelligent algorithms and techniques to optimize agricultural processes and increase crop yields. The project seeks to integrate machine learning capabilities into a sophisticated monitoring and control system that can accurately assess crop growth patterns, predict possible crop diseases and pests, and automate irrigation, fertilization, and harvesting process. The Harvestify Machine Learning project seeks to revolutionize the agriculture industry by leveraging data-driven insights and advanced technologies to enhance productivity, profitability, and sustainability.

# Project Description

The proposed system for crop and fertilizer recommendation and disease prediction using machine learning and deep learning in agriculture provides farmers the details information and it acts as a helping hand. This system acts as a suggestion box for the farmers as suggesting farmers about farming at every stage of crop growth so that it will be prevented from disease. The main aim is to provide the details information of crop from the period of cultivation and process of using different fertilizers at different stages of crop to prevent from various diseases. In this project with the help of some machine learning and deep learning algorithms and techniques there will be prediction of crop and disease on the basis of climate condition, condition, soil type, rainfall and its quality. Using their will be prediction of fertilizer for good crop production. The crop disease detection is done by using some classification algorithms such as segmentation algorithm i.e., CNN.

## Scope of the Work

The scope of work of a crop, fertilizer, and disease prediction model is to provide farmers with information about the best time to plant, fertilize, and harvest their crops, as well as to predict the likelihood of crop diseases. This information can help farmers to improve their crop yields and reduce their risk of crop losses.

The model is typically developed using a combination of historical data, weather data, and crop growth models. The historical data is used to train the model to identify patterns that can be used to predict future crop yields and disease risks. The weather data is used to update the model as conditions change. The crop growth models are used to simulate the growth of crops under different conditions.

The model can be used by farmers to make decisions about their crops. For example, the model can be used to determine the best time to plant a crop based on the expected weather conditions. The model can also be used to determine the amount of fertilizer that should be applied to a crop based on the soil conditions and the expected crop yield. Finally, the model can be used to predict the likelihood of crop diseases so that farmers can take steps to prevent or control them.

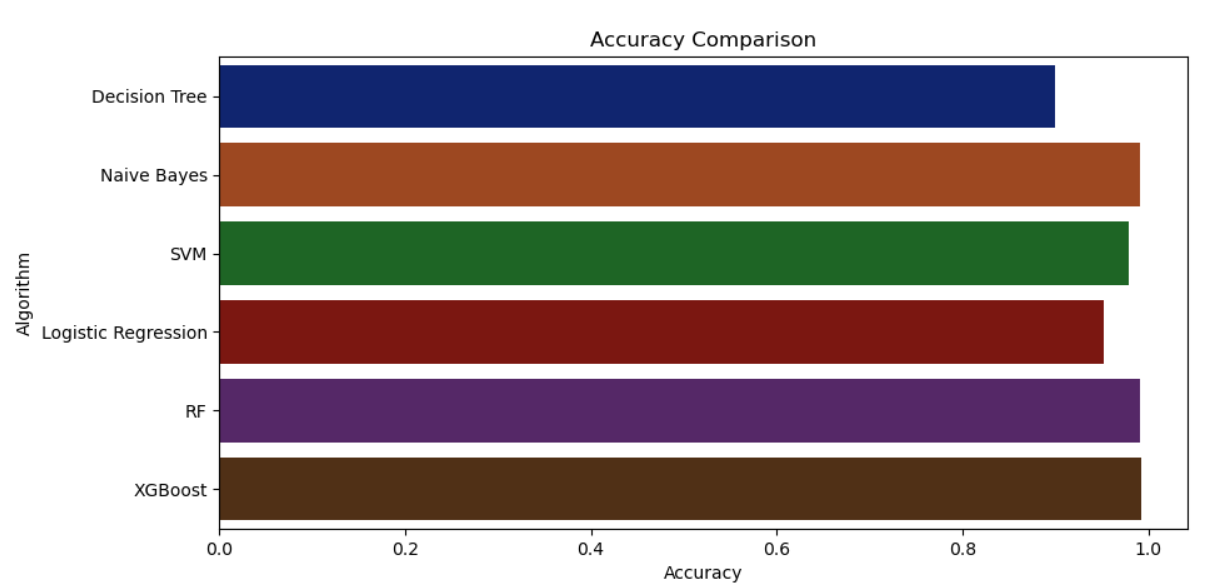
The use of a crop, fertilizer, and disease prediction model can help farmers to improve their crop yields and reduce their risk of crop losses. This can lead to increased profits for farmers and a more secure food supply for the world.

The scope of work for a crop, fertilizer, and disease prediction model can vary depending on the specific needs of the farmers who will be using the model. However, the general tasks that are included in the scope of work are as follows: data collection, data cleaning and preparation, model development, model testing, and model deployment.

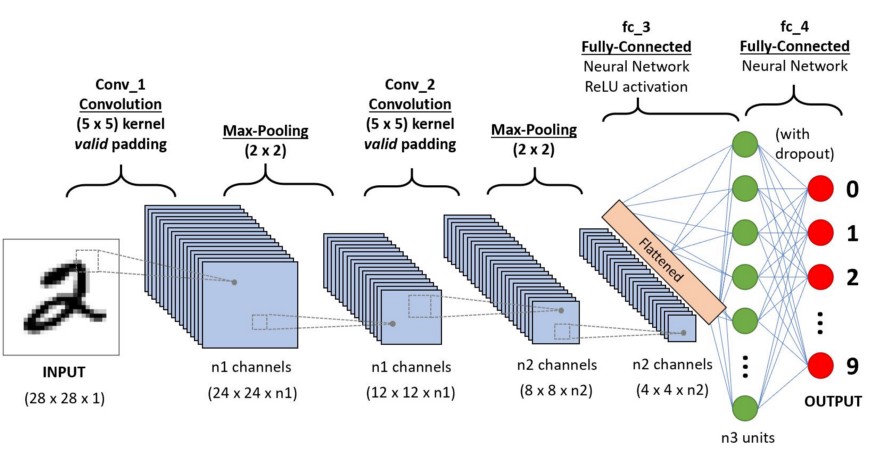
## Project Modules

The project module of crop, fertilizer, and disease prediction model can be divided into the following steps:

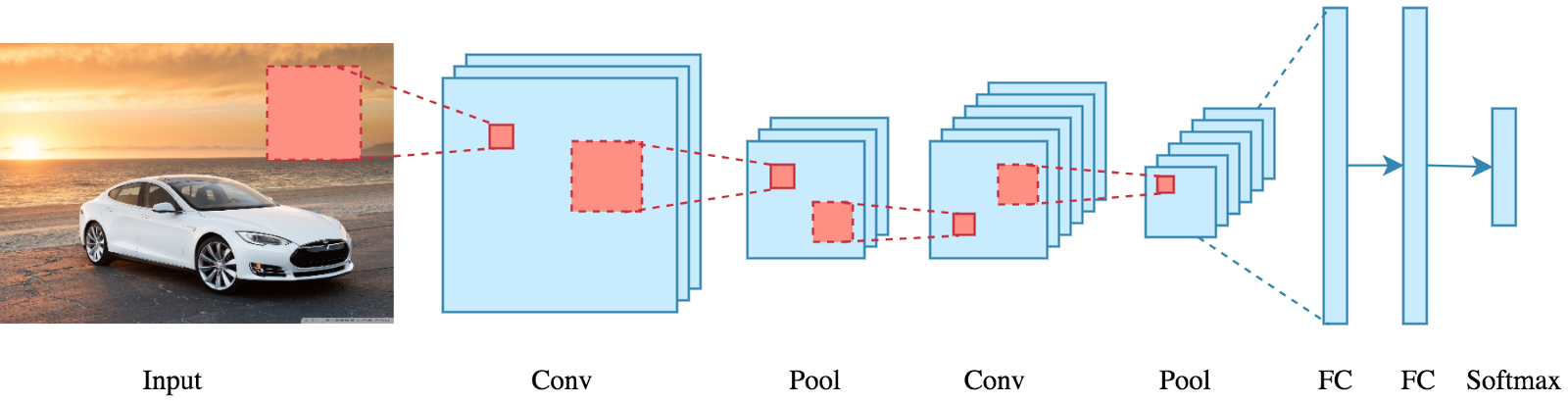
* **Data Module: -** It is the very first module of this project. To train machine learning and deep learning model it is very important to collect necessary dataset. In this module the data collection and data processing has been done. With the help of pre-processing, It select the feature and find the correlation between the attributes of the dataset. In this first we download the dataset from the Internet for crop, fertilizer and disease prediction and then load the dataset into the code and look how and what actually data is. And if there is a NULL value and while cleaning we will replace blanks and NULL values with most occurring value in that column.
* **Feature Engineering: -** This module is also called as dimensionality reduction. The main motive of this model is used remove the redundant features or the error from the give dataset. In feature engineering we create new feature which is helpful in future for outlier detection and removal in later stage. In our dataset label (crop type) is the categorial feature which cause problem if we have so many labels (crop type) at the same time. It creates problem in searching the best crop, fertilizer and disease analysis.
* **Outlier Detection: -** This module is also called as dimensionality reduction. The main motive of this model is used remove the redundant features or the error from the give dataset. In feature engineering we create new feature which is helpful in future for outlier detection and removal in later stage. In our dataset label (crop type) is the categorial feature which cause problem if we have so many labels (crop type) at the same time. It creates problem in searching the best crop, fertilizer and disease analysis.
* **Model Module: -** In this module, firstly we train different-different machine learning algorithm model on the same data and check their accuracy and select the best fit algorithm which give high accuracy result and prediction as a crop recommendation algorithm and use for further process. For Fertilizer Recommendation we use the same algorithm which is used by crop recommendation as they need same data as an input like Nitrogen, Potassium, Phosphorus, Ph value, rainfall in mm, location to predict the crop and fertilizer. For Disease Prediction which is a part of image classification we use the CNN and RS-Net to process the image and predict the result for this we use GPU to train our model for this GPU is not provided on every platform to train the model again and again so we train the model only once’s and save that model for the future uses.   
  The algorithm we check for best accuracy: -
  + **Decision Tree Classifier: -** A decision tree classifier is a supervised learning algorithm that can be used to classify data into two or more categories. It works by recursively partitioning the data into smaller and smaller subsets until each subset contains only data of a single category. The decision tree is built by starting at the root node and asking a question about one of the features of the data. The data is then split into two subsets based on the answer to the question. This process is repeated for each child node until all of the data has been classified.
  + **Support Vector Machine: -** A support vector machine (SVM) is a supervised machine learning algorithm that can be used for both classification and regression tasks. It works by finding a hyperplane in the feature space that separates the data points of two or more classes. The hyperplane is chosen such that the distance between the hyperplane and the closest data points of each class is maximized. This distance is called the margin. SVMs are known for their accuracy and robustness to noise. They are also relatively easy to train and interpret. However, they can be computationally expensive to train, especially for large datasets.
  + **Random Forest Classifier: -** A random forest classifier is an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees. Random forests are a powerful and versatile machine learning algorithm that can be used for a variety of tasks. They are particularly well-suited for tasks where accuracy and robustness are important.
  + **Logistic Regression: -** Logistic regression is a statistical model that is used to predict the probability of an event occurring. It is a type of regression analysis that uses a logistic function to model the probability of a binary outcome. The logistic function is a sigmoid function that maps the real number line to the interval [0, 1]. This means that the logistic function can be used to model probabilities, which are values between 0 and 1. Logistic regression is a popular machine learning algorithm for classification problems. It is used to predict the probability of a categorical outcome, such as whether or not a customer will click on an ad, whether or not a patient has a disease, or whether or not a student will pass an exam. Logistic regression works by fitting a model to a set of data points. The model consists of a linear combination of the features and a bias term. The linear combination is multiplied by a logistic function to produce the probability of the outcome. The model is then used to predict the probability of the outcome for new data points. Logistic regression is a powerful tool for classification problems. It is easy to understand and interpret, and it can be used to handle both categorical and numerical data. However, logistic regression can be sensitive to overfitting, so it is important to use appropriate techniques to prevent this.
  + **Gaussian Naïve Bayes: -** Gaussian naive Bayes is a supervised machine learning algorithm that is used for classification problems. It is based on the Bayes theorem, which states that the probability of an event A happening, given that event B has already happened, is equal to the probability of event A happening times the probability of event B happening given that event A has already happened, divided by the probability of event B happening. Gaussian naive Bayes assumes that the features of the data are normally distributed. This means that the data can be described by a mean and a standard deviation. The mean is the average value of the data, and the standard deviation is a measure of how spread out the data is. Gaussian naive Bayes works by first calculating the mean and standard deviation of each feature for each class. Then, it uses the Bayes theorem to calculate the probability of each class given the data. The class with the highest probability is the class that the data is predicted to belong to. Gaussian naive Bayes is a simple and efficient algorithm that can be used for a variety of classification problems. It is particularly well-suited for problems where the features are normally distributed.
  + **XGBoost: -** XGBoost, also known as Extreme Gradient Boosting, is an open-source software library for machine learning that uses the gradient boosting algorithm. It is used for classification, regression, and ranking problems. XGBoost is a popular choice for machine learning competitions because it is known to achieve state-of-the-art results. XGBoost is a decision tree-based algorithm that works by iteratively adding new trees to the model. Each new tree is added to correct the errors made by the previous trees. This process is called boosting. XGBoost uses a variety of techniques to improve the accuracy and performance of the model. XGBoost is a powerful and versatile machine learning algorithm that can be used for a variety of problems. It is particularly well-suited for problems where accuracy and performance are important.



* + **CNN Model: -** CNNs were first developed and used around the 1980s. The most that a CNN could do at that time was recognize handwritten digits. It was mostly used in the postal sectors to read zip codes, pin codes, etc. The important thing to remember about any deep learning model is that it requires a large amount of data to train and also requires a lot of computing resources. This was a major drawback for CNNs at that period and hence CNNs were only limited to the postal sectors and it failed to enter the world of machine learning.

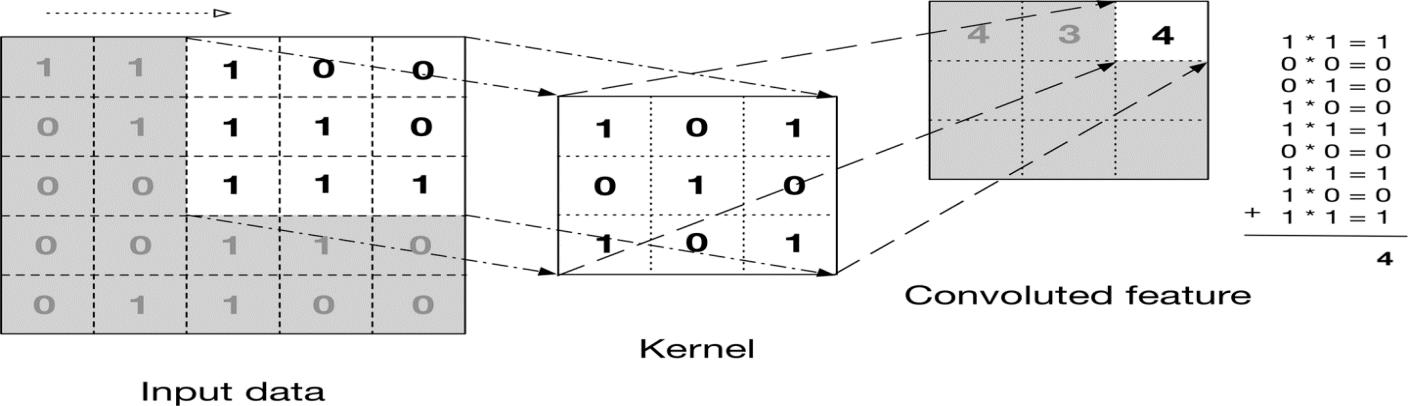


In deep learning, a convolutional neural network (CNN/ConvNet) is a class of deep neural networks, most commonly applied to analyses visual imagery. Now when we think of a neural network, we think about matrix multiplications but that is not the case with ConvNet. It uses a special technique called Convolution. Now in mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other.

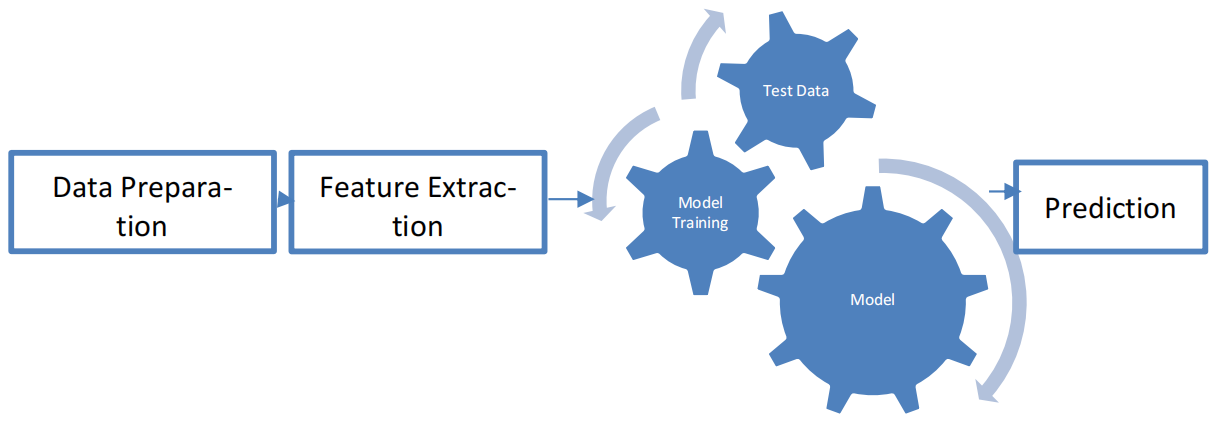


But we don’t really need to go behind the mathematics part to understand what a CNN is or how it works.

Bottom line is that the role of the ConvNet is to reduce the images into a form that is easier to process, without losing features that are critical for getting a good prediction.



* Application Module: - This module is divided into three parts i.e., Crop Recommendation, Fertilizer Recommendation and Disease Prediction. For crop recommendation we take several inputs of soil testing report and taking location as an input and through weather API we fetch the weather report of that location such as rainfall and weather report. For fertilizer recommendation we take soil report and the crop and based upon it it will recommend the fertilizer. For disease prediction model it will take image as an input and then classify the result and predict the disease and crop and give prevention method and solutions.



The choice of algorithm will depend on the specific problem that is being addressed. For example, decision trees are often used for classification problems, while neural networks are often used for regression problems.

The crop, fertilizer, and disease prediction model can be used to help farmers make better decisions about their crops. For example, the model can be used to recommend the best crops to grow in a particular region, based on the soil conditions and climate. The model can also be used to recommend the best fertilizers to use for a particular crop, and to detect and diagnose plant diseases.

The use of machine learning to predict crop, fertilizer, and disease can help farmers to improve their yields and reduce their costs. This can lead to increased profits for farmers and improved food security for the world.

## Context Diagram (High Level)

A zero-level data flow diagram (DFD) is the highest level of a DFD. It shows the overall system or process being analyzed or modeled. It is designed to be an at-a-glance view, showing the system as a single high-level process, with its relationship to external entities. It should be easily understood by a wide audience, including stakeholders, business analysts, data analysts and developers. The process symbol represents the overall system or process being analyzed or modeled. The external entity symbols represent the people, organizations, or other systems that interact with the system. The data flow lines represent the flow of data between the process and external entity symbols. Zero-level DFDs are a useful tool for understanding the overall structure of a system. They can be used to identify the key components of the system and the relationships between them.



# Implementation Methodology

Data Collection

Data Preprocessing

Remove Outliers

Training Dataset

Jupyter Lab

Decision Tree Classifier

Random Forest

Support Vector Machine

XGboost

Gaussian Naïve Bayes

CNN and RS-Net

Model Building

Predicted Value

The main aim of this project to be implemented is to recommend best fit crop and fertilizer and predict the disease via image classification and give the prevention and solution method. Here step by step process involved is represented below.

1. Scientific Environment
2. Source of Data
3. Excel 2016: The first process to store the data
4. Loading data into notebook.
5. Normalizing of data
6. Detecting outliers
7. Machine Learning model are built, and various algorithm used for predicting as listed in methodologies.
8. Split the data sets as testing and train data for cross validator process.
9. Fitting the data into different machine learning for prediction.
10. Finding the Error to find the accuracy percentage.
11. Visualization: Making a web page and deploy it so that everyone can use it and for user friendly environment.

* **Need of Technical Enviroment**
  + Microsoft Excel.
  + Jupyter Notebook/Google Collab for creating the Script.
  + Essential Libraries to be installed.
  + Regressing and Classification
  + Appropriate Function to be selected for algorithm
  + Web Designing.
* **Data Sources**

As I mentioned before the main data sets were collected from the Kaggle machine learning repository which is open data resource for the data mining and for the predictive analytics purposes. The acquired data source was a text file for finding the errors in the data source, the text file is connected to the EXCEL and they are separated using the commas and saved as a CSV file.

* **Data Cleaning**

The data in the CSV files need to be checked whether it has any missing values, as the data source have missing values every attribute have checked using the filters and null values are removed which helps to increase the accuracy level.

* **Tool 1 Jupyter lab/Notbook**

The Crop and fertilizer data is processed with the help of python program in Jupyter Notebook and the numerical data is normalized using the script so the data will equally distributed and variables cannot dominate each other. It helps to bring the values of the attribute on a common scale.

* **Implementation**

The data is been split into test and train 75 percent data is used for the test data and remaining for the train data. Consider the standard recursive partitioning algorithm will start it searches all the data and in-depth search is made for all the variables and the best split point is taken and the process gets repeated for the right and left leaves. The method has been followed for different assumption here the dependent variable is values of the soil and the location and attribute which is going to be predicted and the other attributes are independent variables and the attributes are given into the model and with the help of coefficients an equation has been made for regression and the predicted prices are produced.

# Technologies to be used

## Software Platform

1. **Front-end**
   * + 1. HTML
       2. CSS
       3. JS
       4. BootStrap
       5. Font Awesome
       6. Glyphicon
2. **Back-end**
   * + 1. Python
       2. Flask
       3. Jinja Script
       4. TensorFlow
       5. Machine Learning Model
       6. Deep Learning Model
       7. Weather API
       8. Torch

## Hardware Platform

* RAM
* Hard Disk
* OS
* Browser
* Graphic Card
* Powerful CPU/GPU, etc.

## Tools, if any

|  |  |  |
| --- | --- | --- |
| **Name** | **Version** | **Purpose** |
| Python | 3.7 | Primary Programming Languages |
| Google Collab | 1.0.0 | Editor or IDLE |
| JavaScript | 13th Edition | For Making responsive |
| Sklearn | .21.0 | Machine Learning Libraries |
| Matplotlib | 3.0 | Used for graph and plotting |
| Pandas | 1.3.5 | Data Analysis Tool |
| Bootstrap | 5.2 | Web Development tool |
| Pytorch | 1.21.1 | Used for Deep learning neural network |
| TourchVision | 0.13.1 | Neural Network |
| Pickle | 3.5 | For Saving Model in file/data structure |

# Advantages of this Project

There are several advantages to using machine learning in the Harvestify project:

1. Better Yield Predictions: Machine learning algorithms can analyze data on crop yields, weather patterns, soil types, and other environmental factors to predict with high accuracy, the crop yield. This helps farmers make informed decisions regarding crop rotation, fertilization and irrigation, which ultimately leads to an increase in the overall productivity of crops.
2. Efficient Resource Management: Machine learning algorithms can identify factors that are critical for crop growth, such as temperature, humidity, sunlight and soil moisture, allowing farmers to optimize the resources needed for each crop. This helps to reduce waste and save resources, making farming more efficient and cost-effective.
3. Improved Precision Agriculture: With the help of machine learning, farmers can implement precision agriculture practices. This involves using sensors and drones to monitor the health and growth of crops in real-time. Machine learning algorithms can process this data to identify any issues with crops, allowing farmers to intervene early to protect their crops.
4. Better Decision-making: Machine learning algorithms can analyze large datasets on weather patterns, soil types, and crop yields, enabling farmers to make well-informed decisions. They can be used to evaluate the impact of various factors on the agricultural yield, allowing farmers to optimize their decisions.
5. Higher Efficiency: The use of machine learning algorithms in the Harvestify project helps to significantly reduce the time and effort required to monitor crops, analyze data, and make decisions. This leads to higher productivity, improved yield, and greater profitability for farmers.

# Assumptions, if any

However, based on general assumptions, here are some possible assumptions:  
The project aims to improve crop yield and quality through the use of machine learning algorithms.

1. The project involves collecting and analyzing data from various sources, such as weather patterns, soil conditions, and crop growth stages.
2. The project uses supervised and unsupervised learning techniques to identify patterns and make predictions about crop growth and yield.
3. The project involves the development of a user-friendly interface that allows farmers to input data and receive recommendations based on the machine learning models.
4. The project may require significant computational resources, such as high-performance computing clusters, to process large amounts of data.
5. The project may face challenges related to data privacy and security, as sensitive information about farmers' crops and land may be involved.
6. The success of the project depends on the accuracy of the machine learning models and the adoption of the technology by farmers.

# Future Scope and further enhancement of the Project

The Harvestify project can be enhanced further by incorporating machine learning models to predict crop yields and optimize farming techniques. This can be achieved by using historical data on weather patterns, soil conditions, and crop productivity to develop predictive models. The machine learning models can also be used to identify pest and disease outbreaks, enabling farmers to take corrective actions in a timely manner.

Another possible enhancement of the Harvestify project is to use machine learning to personalize recommendations to farmers based on their individual farm characteristics, soil type, and weather patterns. This personalized approach could optimize farming techniques and help farmers achieve better crop yields.

In addition, the Harvestify project can be further enhanced by incorporating blockchain technology to ensure transparency and traceability in the supply chain. This can help farmers to prove the authenticity and quality of their crops, which can increase customer trust and enable them to command higher prices.

Overall, the future scope and further enhancement of the Harvestify project using machine learning models are vast, and with the right implementation, it can bring significant benefits to farmers and the agriculture sector as a whole.

# Project Repository Location

| **S#** | **Project Artifacts (softcopy)** | **Location** (Mention Lab-ID, Server ID, Folder Name etc.) | **Verified by Project Guide** | **Verified by Lab In-Charge** |
| --- | --- | --- | --- | --- |
|  | Project Synopsis Report (Final Version) | https://github.com/atishayjain072/Harvestify |  |  |
|  | Project Progress updates | https://github.com/atishayjain072/Harvestify |  |  |
|  | Project Requirement specifications | https://github.com/atishayjain072/Harvestify |  |  |
|  | Project Report (Final Version) | https://github.com/atishayjain072/Harvestify |  |  |
|  | Test Repository | https://github.com/atishayjain072/Harvestify |  |  |
|  | Project Source Code (final version) with executable | https://github.com/atishayjain072/Harvestify |  |  |
|  | Any other document | https://github.com/atishayjain072/Harvestify |  |  |

# Definitions, Acronyms, and Abbreviations

|  |  |
| --- | --- |
| **Abbreviation** | **Description** |
| MSIE | Microsoft Internet Explorer |
| DFD | Data Flow Diagram |
| ER Diagram | Entity Relationship Diagram |
| DAC | Department of Agriculture and Cooperation |
| NIC | National Informatics Center |

# Conclusion

The system uses supervised and unsupervised Machine learning algorithms and gives best result based on accuracy. The results of the two algorithms will be compared and the one giving the best and accurate output will be selected. Thus, the system will help reduce the difficulties faced by the farmers and stop them from attempting suicides. It will act as a medium to provide the farmers efficient information required to get high yield and thus maximize profits which in turn will reduce the suicide rates and lessen his difficulties.

The prediction of crop yield based on location and proper implementation of algorithms have proved that the higher crop yield can be achieved. From above work I conclude that for soil classification Random Forest is good with accuracy 86.35% compare to Support Vector Machine. For crop yield prediction Support Vector Machine is good with accuracy 99.47% compare to Random Forest algorithm. The work can be extended further to add following functionality. Web application can be built to help farmers by uploading image of farms. Crop diseases detection using image processing in which user get pesticides based on disease images. Implement Smart Irrigation System for farms to get higher yield.

We have worked on a sample dataset from Kaggle which has taken into consideration records obtained from a broad agricultural demography. Farmers generally use hit and trial method which leads to wastage of land and resources or even disproportionate growth of crops. We are trying to break all such taxing walls by providing them with an accurate and justified model made by machine learning using random forest classifier to identify the correct crop to be grown in their farms. This will help them in improving their crop production both qualitatively and quantitatively. This will also help them to maintain the quality and nutrition contents of the soil.

In conclusion, the Harvestify project utilizing machine learning has shown great potential in revolutionizing the agriculture industry. Through the use of advanced algorithms and models, farmers can now accurately predict crop yields, identify potential diseases or pests, and optimize irrigation and fertilization practices.   
The integration of machine learning technology in agriculture has the potential to increase crop yields, reduce costs, and improve overall efficiency. Additionally, it can also help farmers make more informed decisions by providing them with real-time data and insights.  
However, it is important to note that the success of machine learning in agriculture depends on the quality and quantity of data available. Therefore, it is crucial for farmers to collect and store accurate and relevant data to ensure optimal results.  
Overall, the Harvestify project using machine learning is a promising development in the agriculture industry, and its continued advancement and implementation could lead to significant improvements in food production and sustainability.

# References

* <https://www.youtube.com/>
* <http://cs229.stanford.edu/>
* <https://ieeexplore.ieee.org/abstract/document/8697639>
* <https://linuxhint.com/>
* [https://towardsdatascience.com/](https://towardsdatascience.com/predicting-house-prices-with-linear-regression-machine-learning-from-scratch-part-ii-47a0238aeac1)
* [https://www.kaggle.com/](https://www.kaggle.com/code/ashydv/housing-price-prediction-linear-regression)
* <https://studygyaan.com/data-science-ml/>
* <https://www.jetir.org/papers/JETIR2110302.pdf>
* Bhagat, N., Mohokar, A., & Mane, S. (2016). International Journal of Computer Applications, 152(2), 23–26.
* N. N. Ghosalkar and S. N. Dhage,” 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1-5, Doi: 10.1109/ICCUBEA.2018.8697639.
* C. R. Madhuri, G. Anuradha, and M. V. Pujitha, A Comparative Study,” 2019 International Conference on Smart Structures and Systems (ICSSS), Chennai, India, 2019, pp. 1-5, Doi: 10.1109/ICSSS.2019.8882834
* T. D. Phan: The Case of Melbourne City, Australia,” 2018 International Conference on Machine Learning and Data Engineering (iCMLDE), Sydney, Australia, 2018, pp. 35-42, Doi: 10.1109/iCMLDE.2018.00017.
* Asia: Survey. (2020). Press Trust of India. https://www.business-standard.com

**Annexure A**

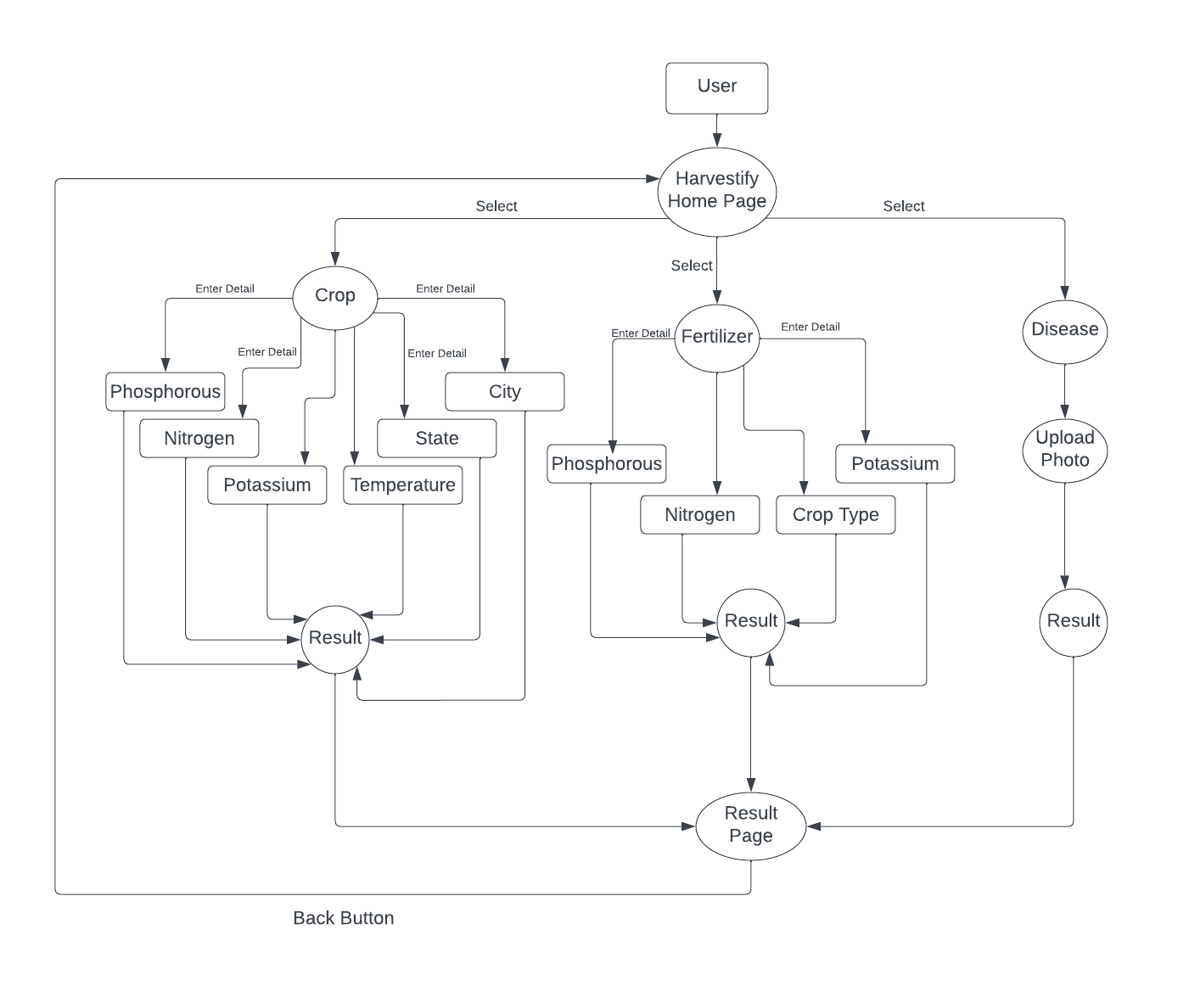
**Data Flow Diagram (DFD)**

**(Mandatory)**

* **Zero Level DFD**



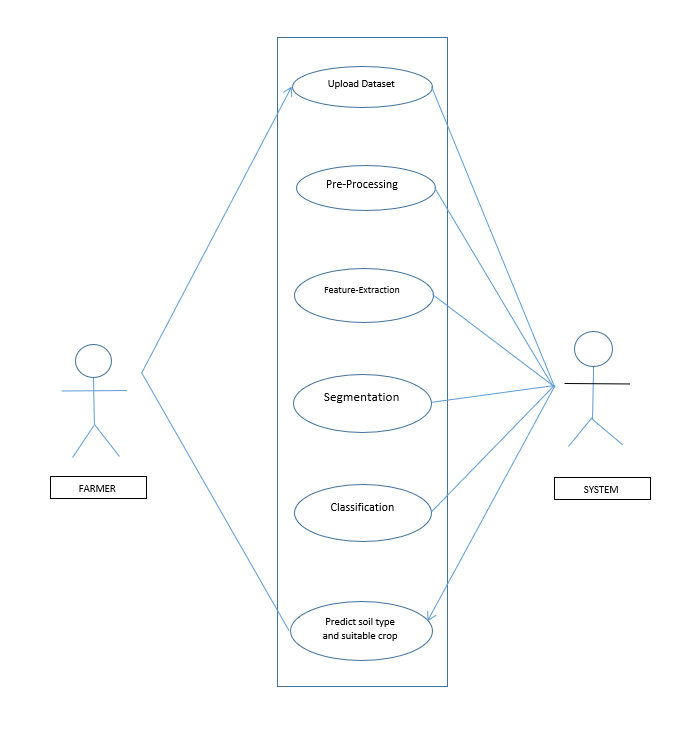
**Data Flow Diagram (First Level)**



**Annexure B**

**Use-Case Diagram (UCD)**

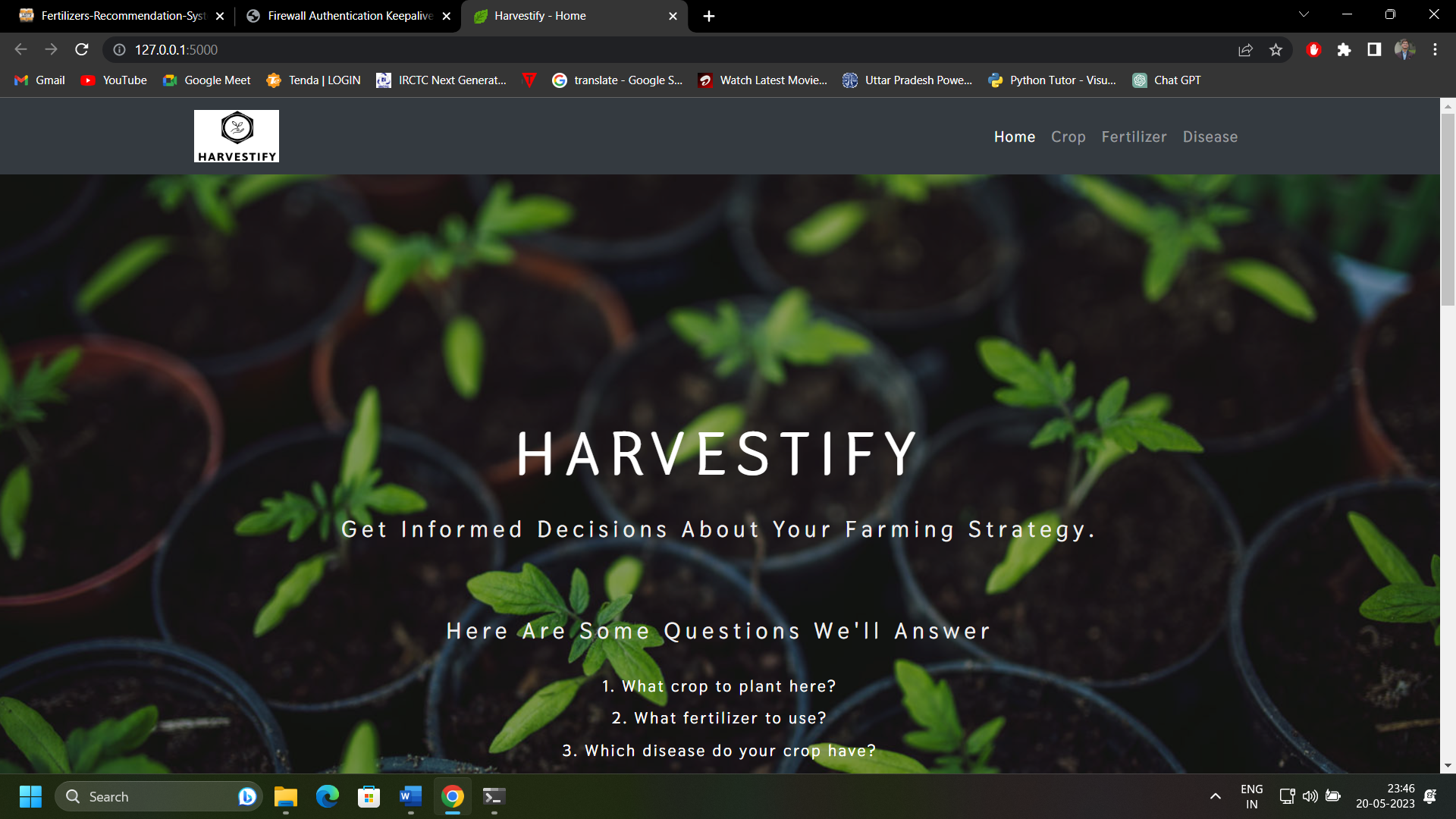
**(Optional)**

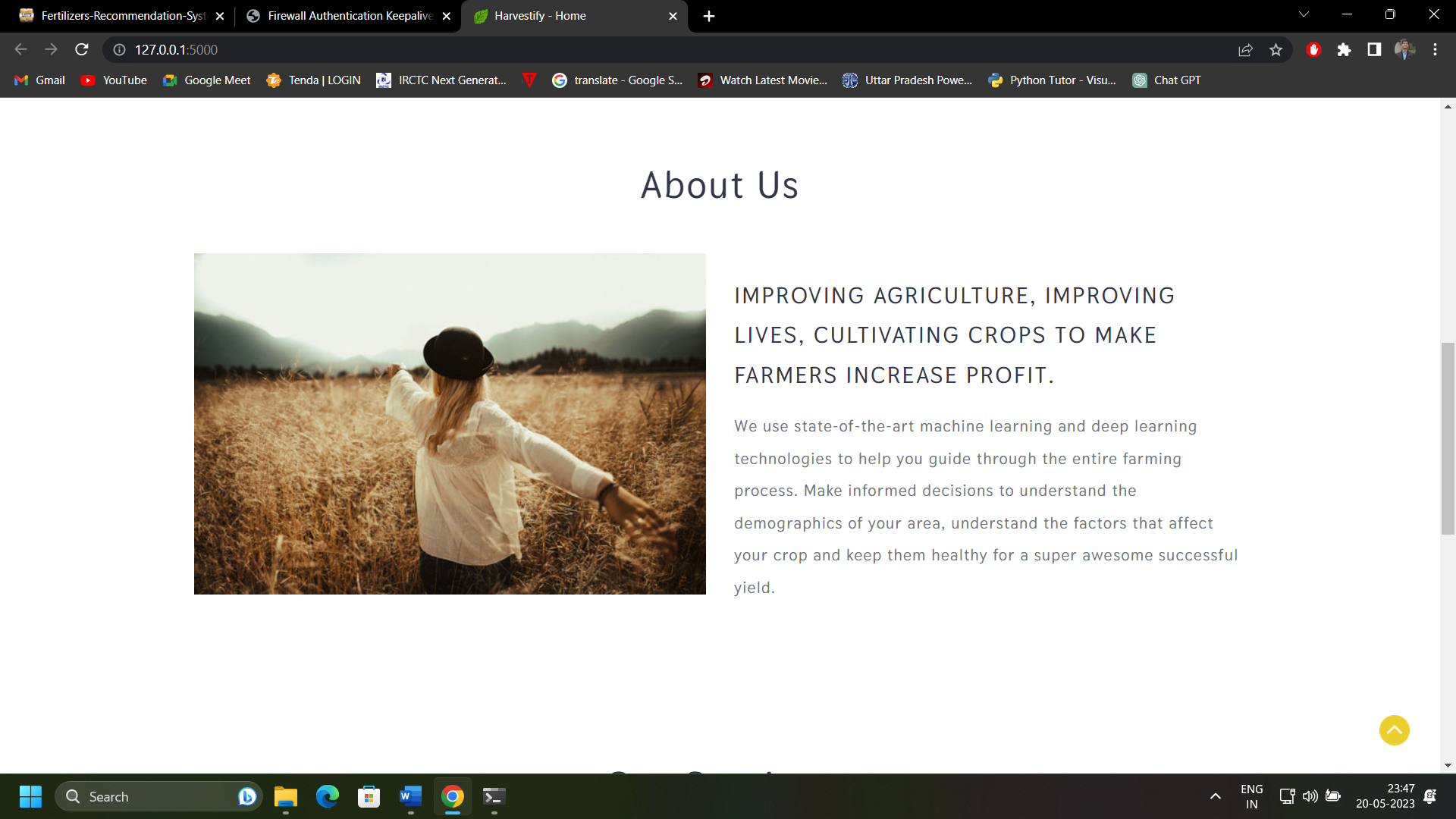
****

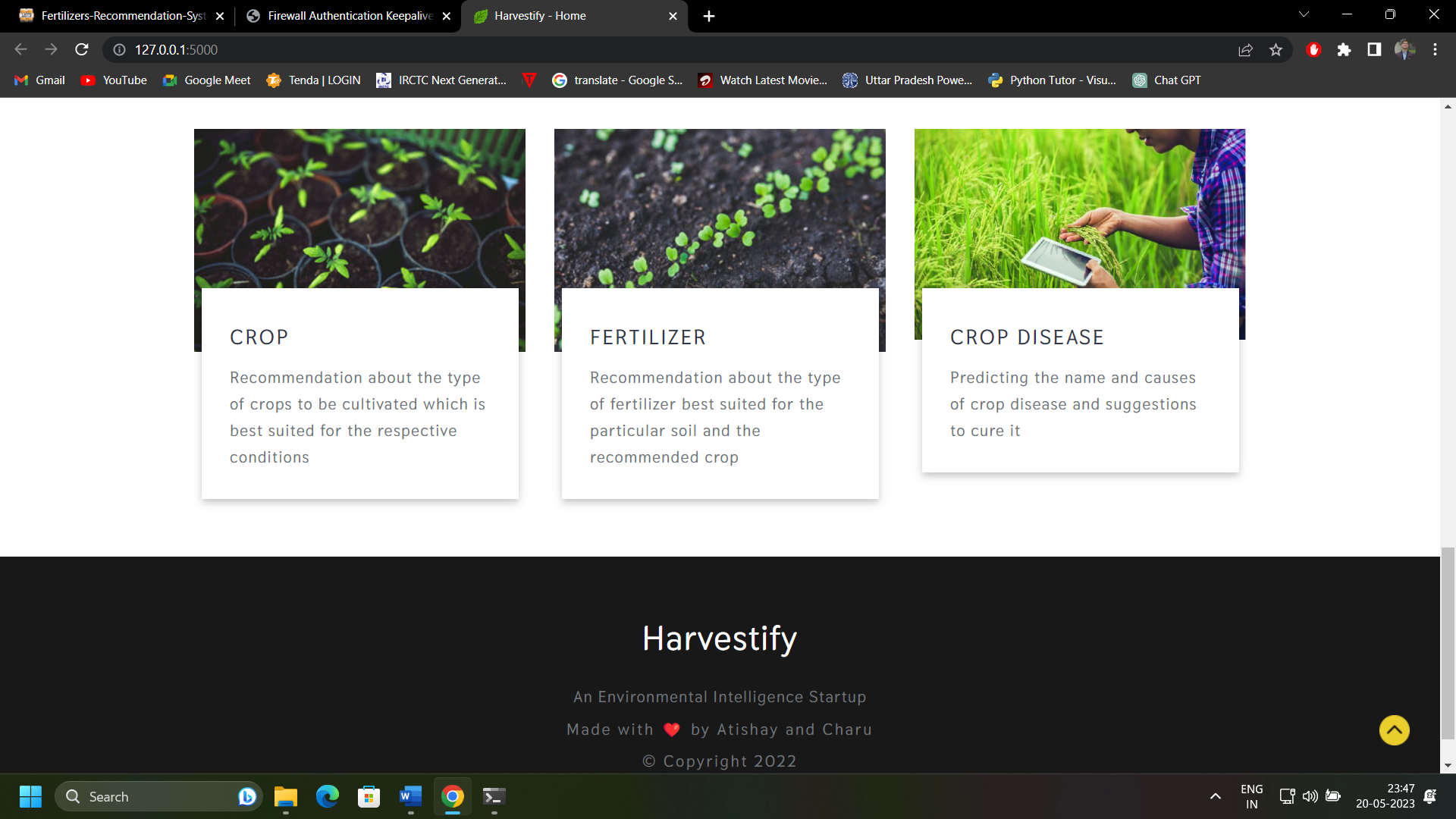
**Annexure C**

**Screen Shots**

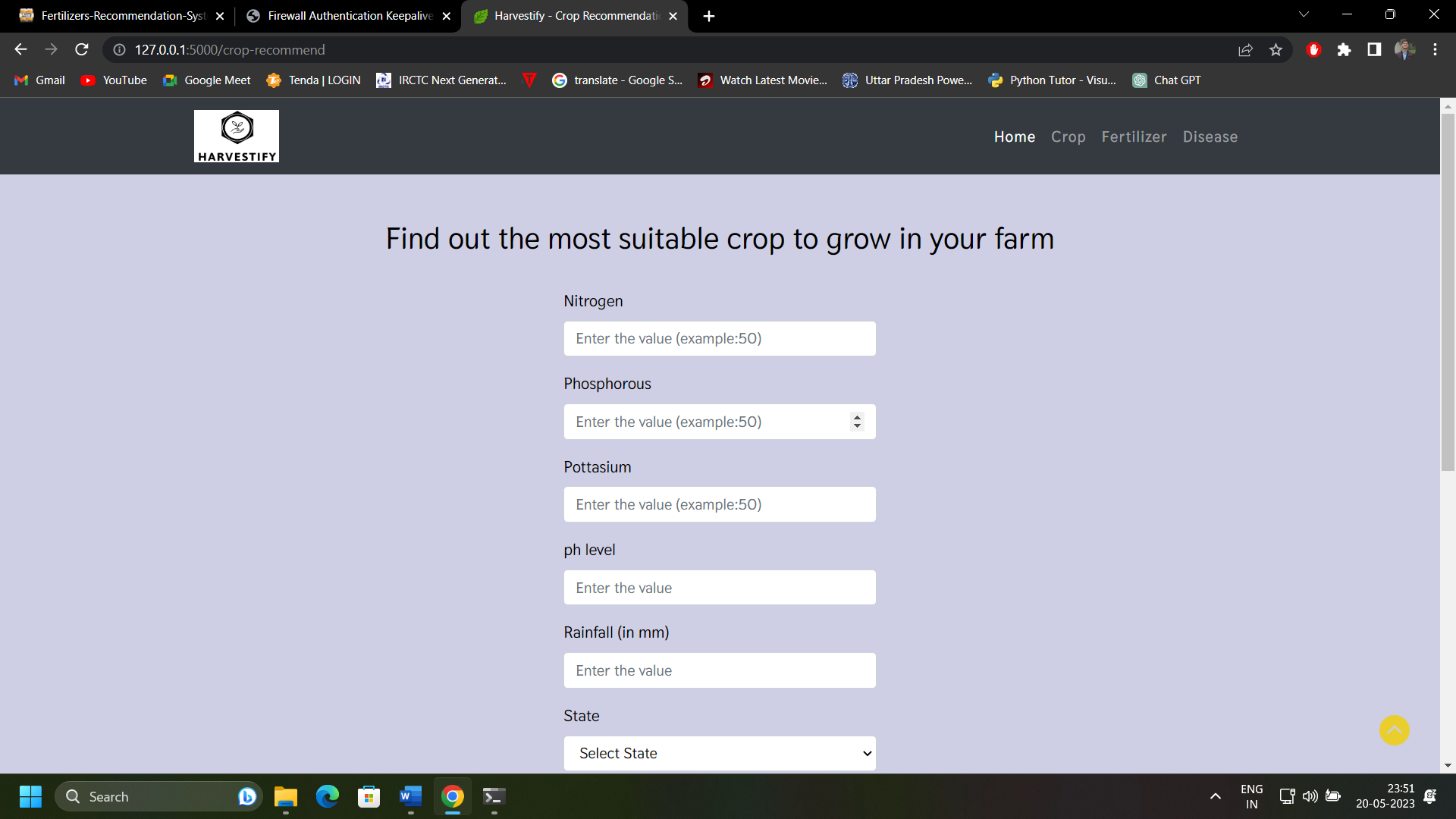
**HOME PAGE:**

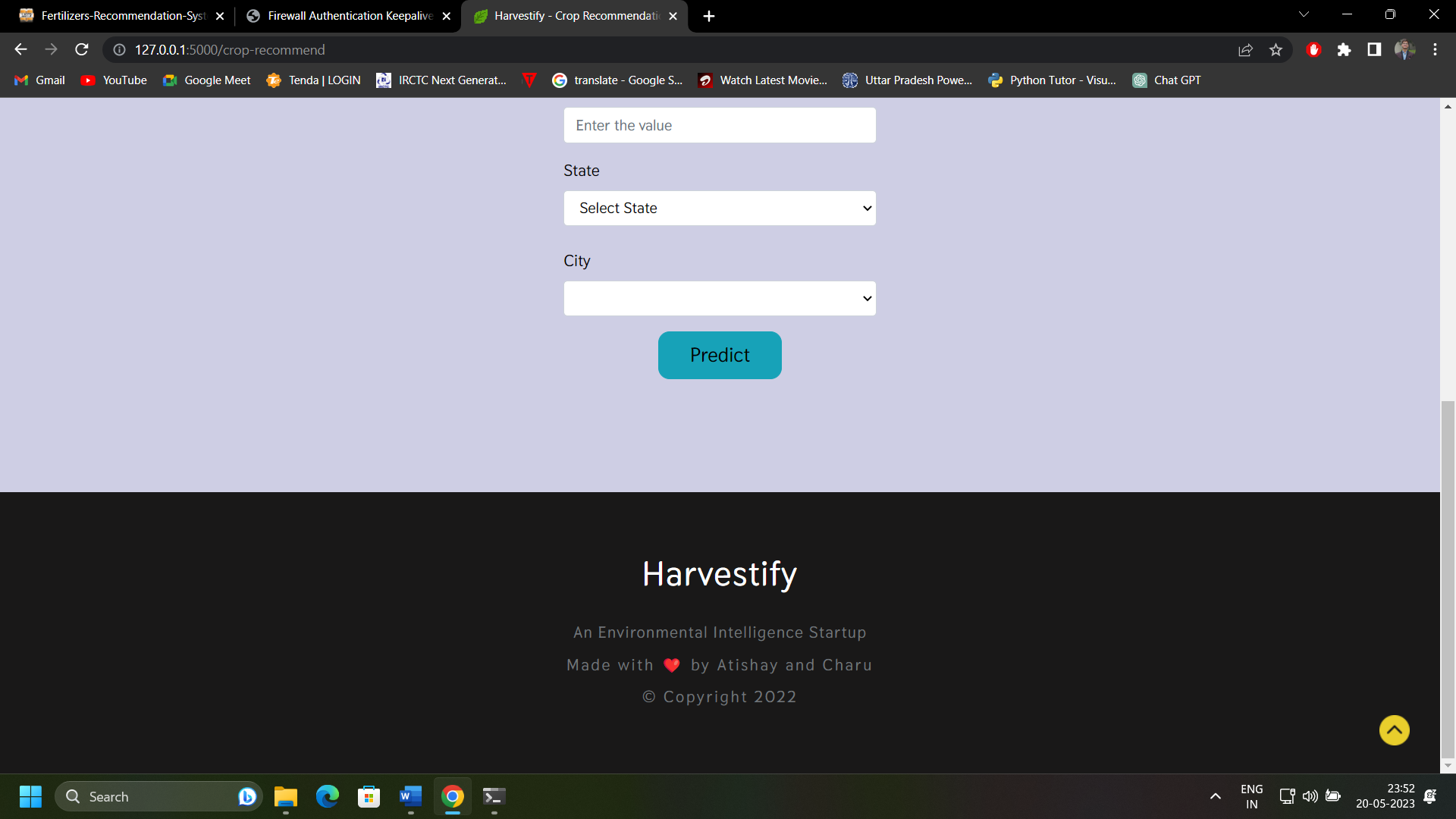
****

****

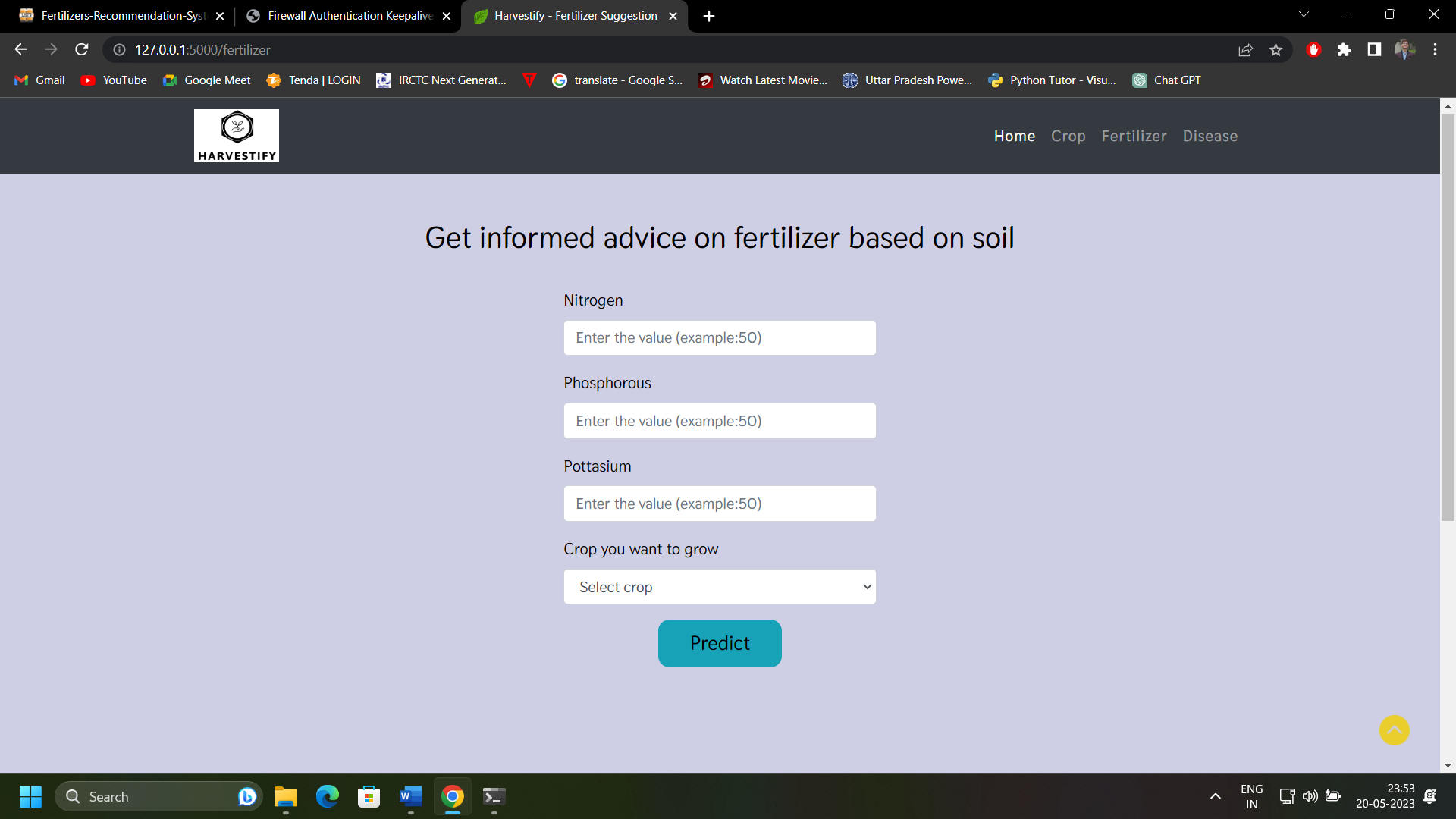
****

**CROP PAGE:**

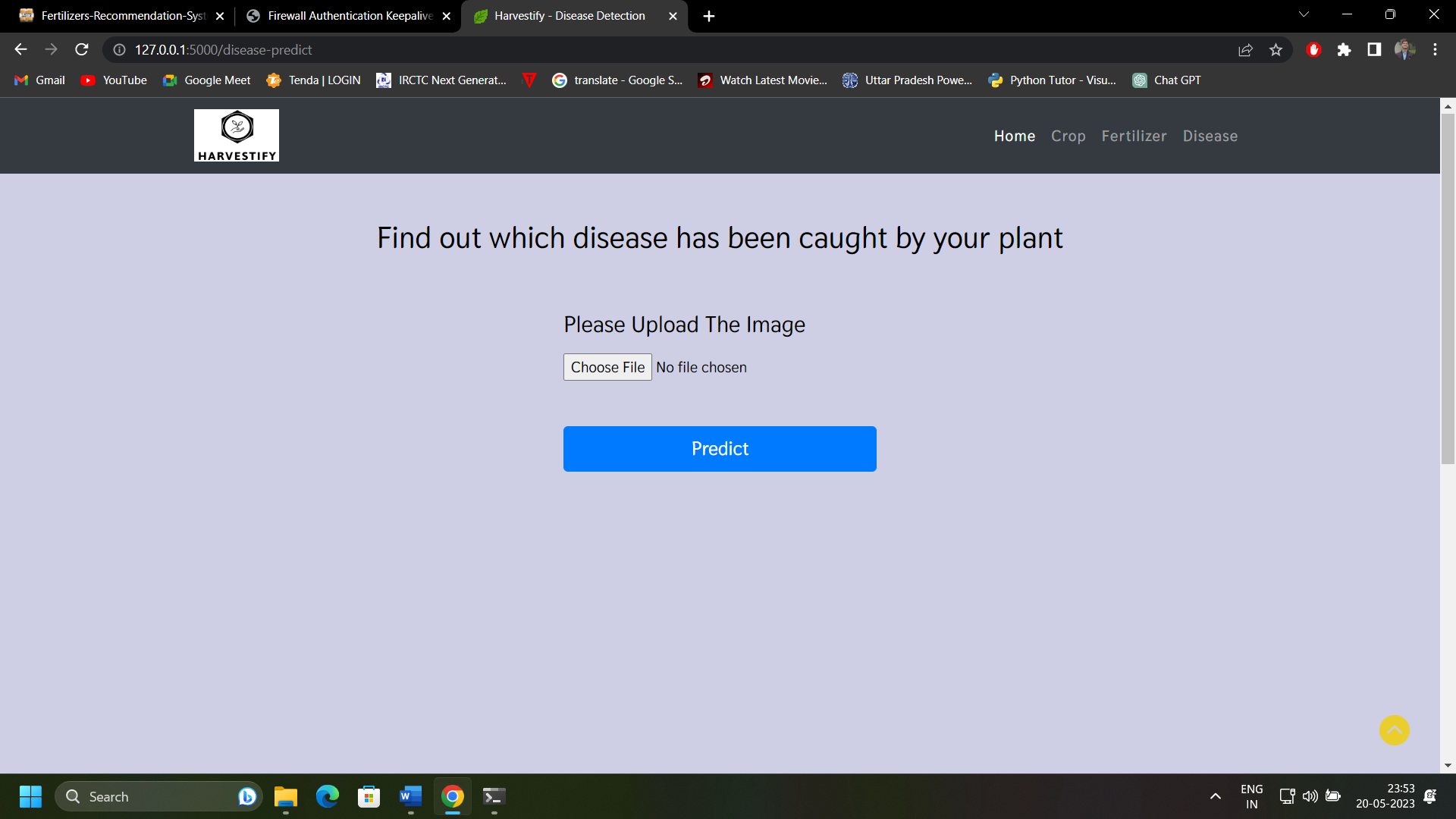
****

****

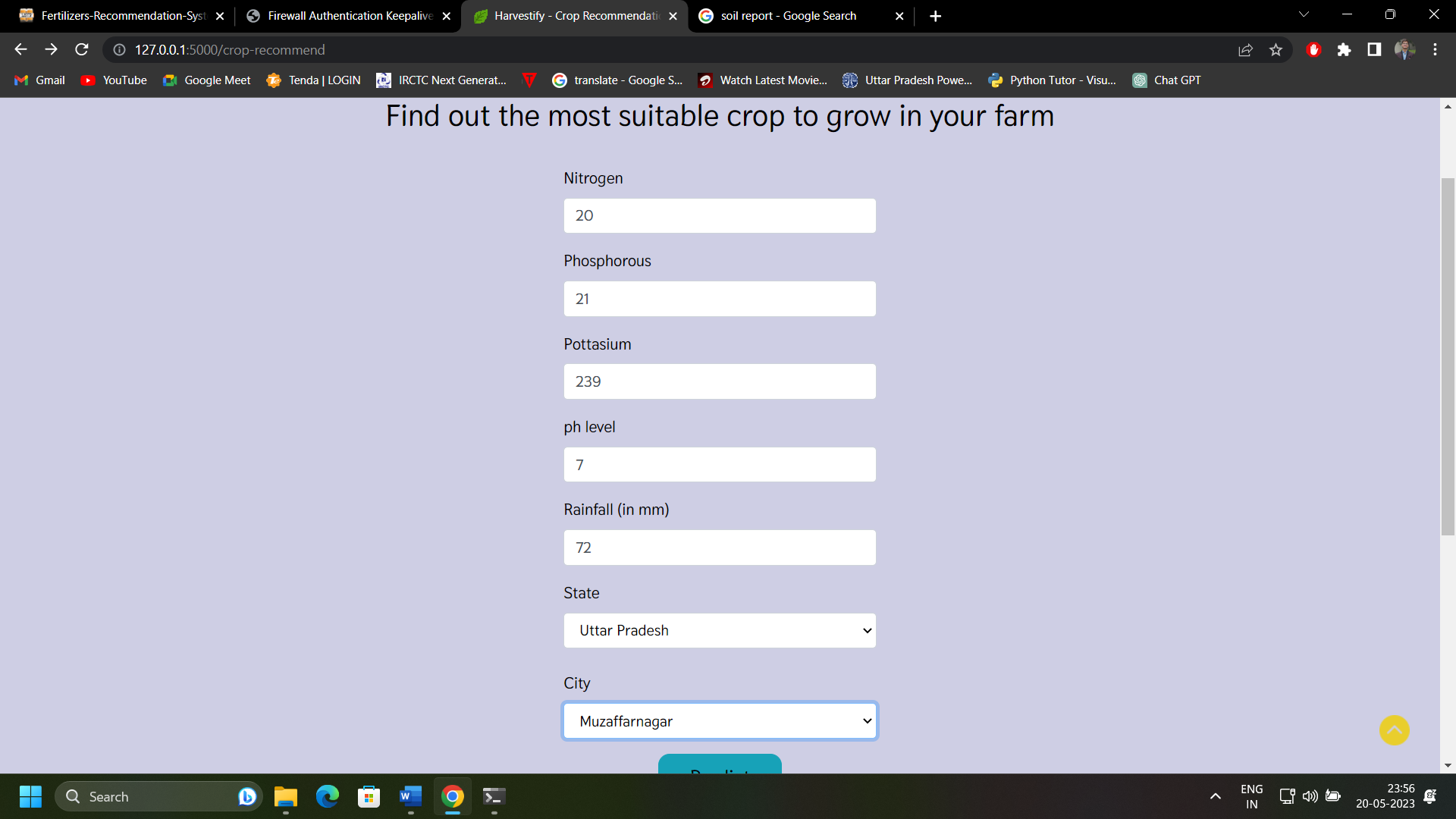
**FERTILIZER PAGE:**

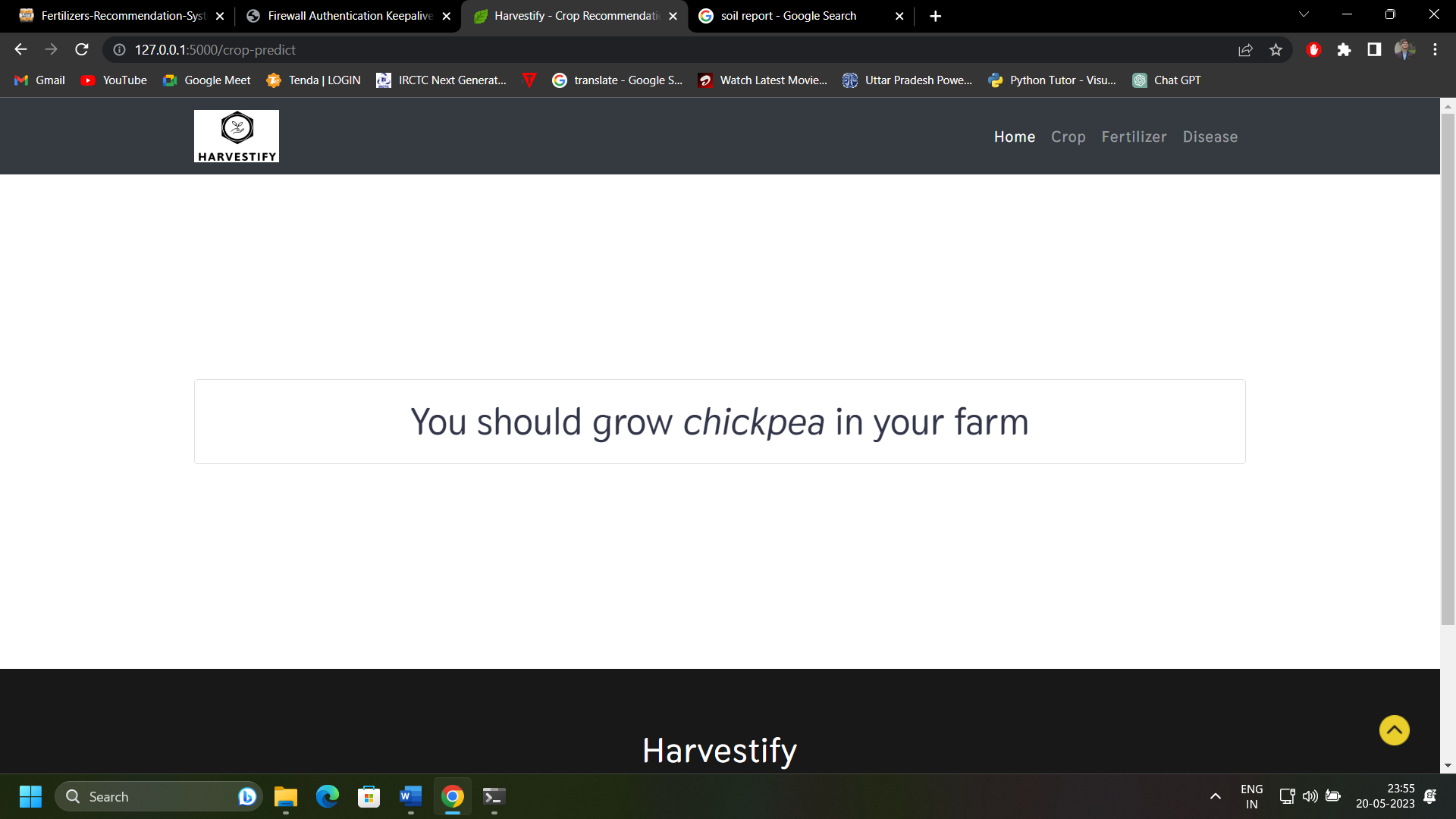
****

**DISEASE PAGE:**

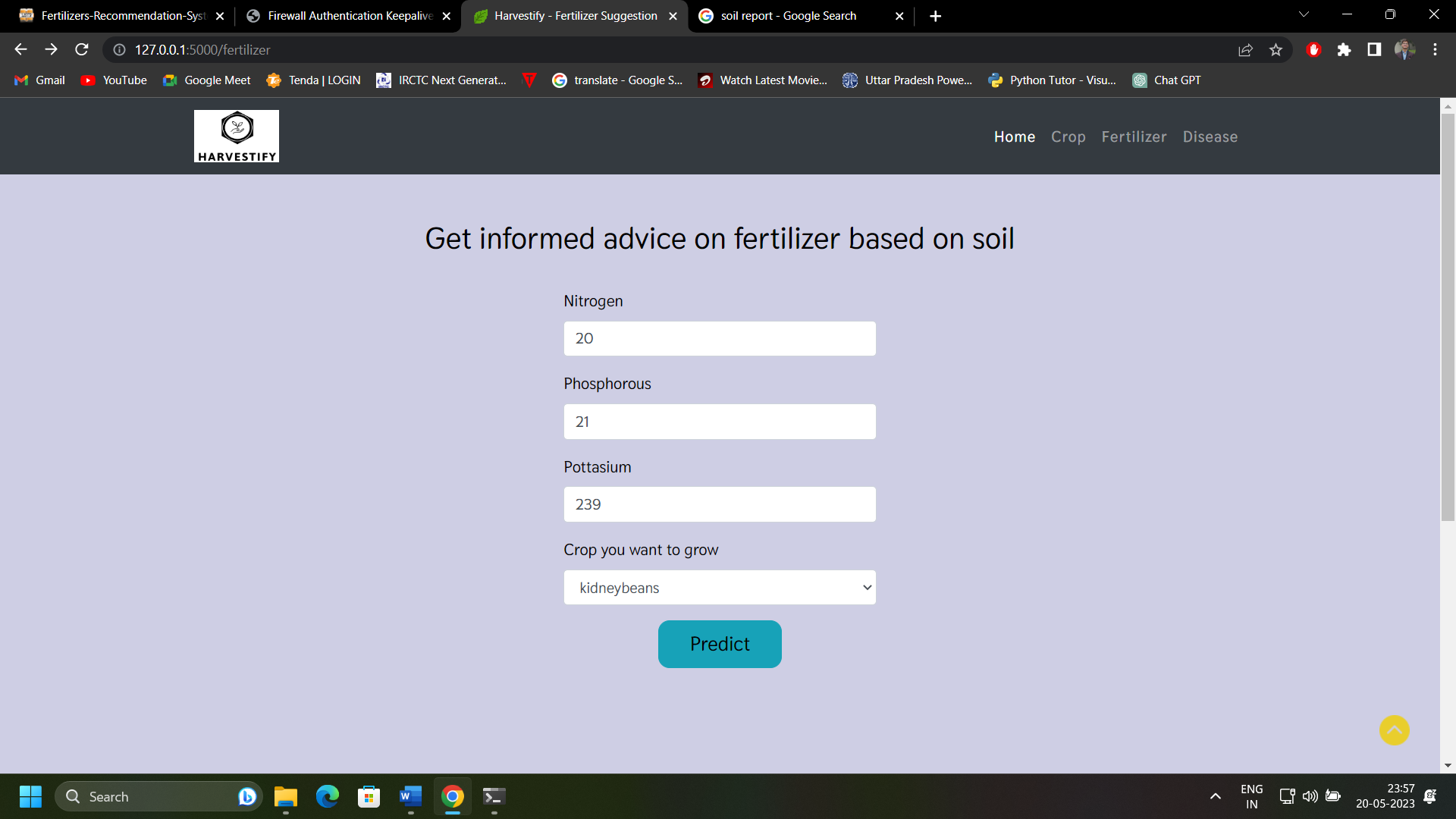


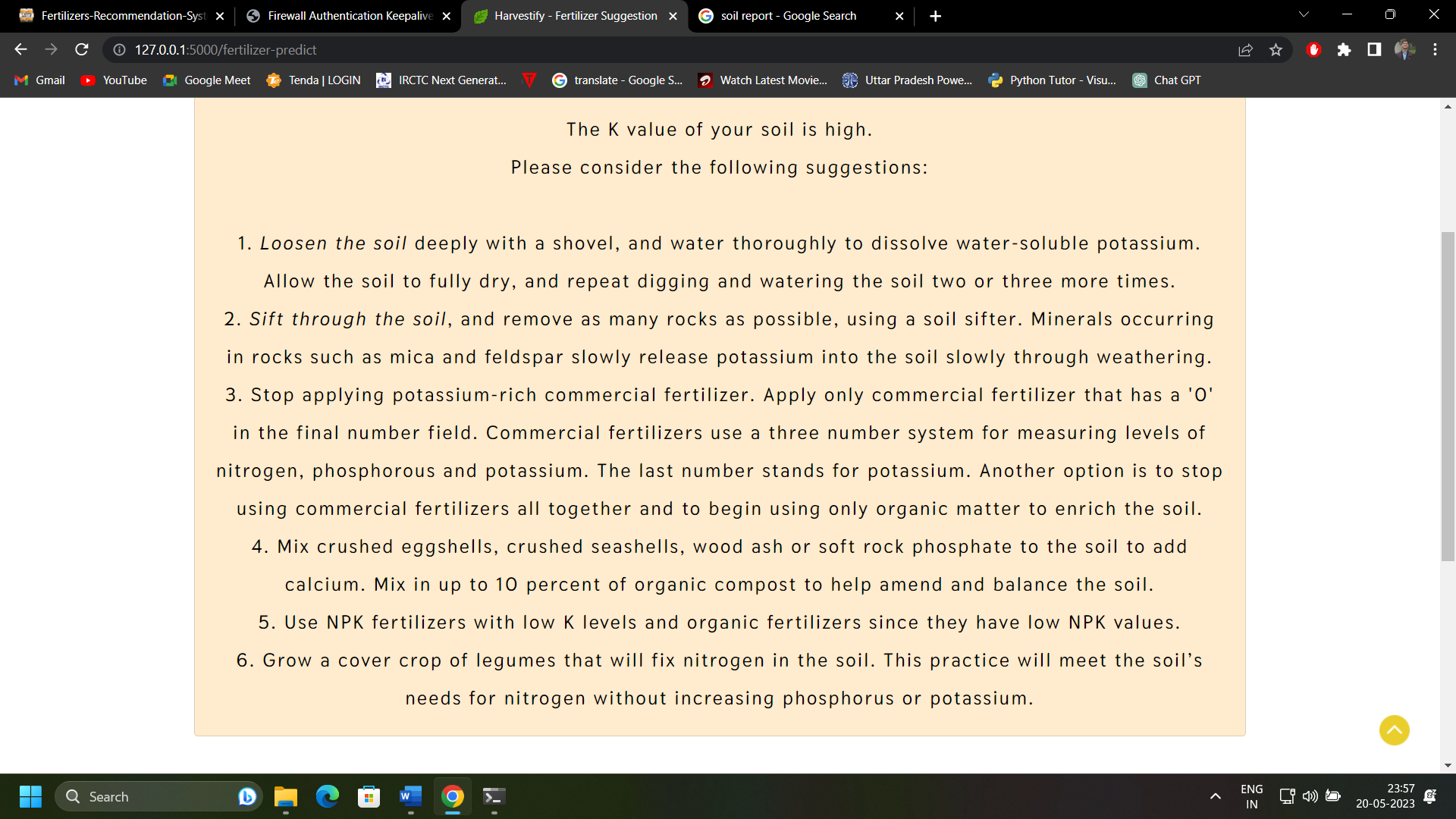
**CROP PREDICTION RESULT:**

****

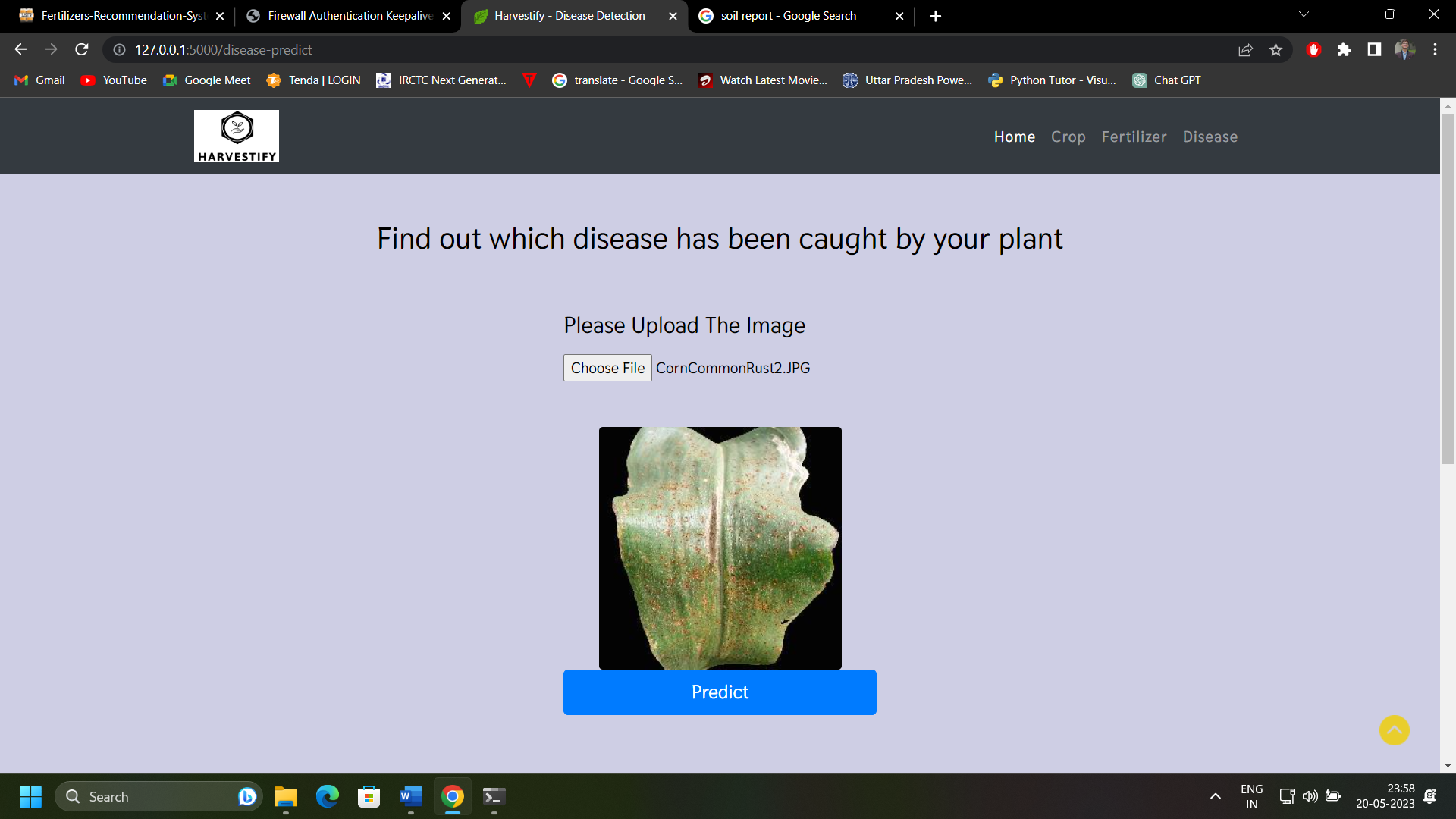
****

**FERTILIZER PREDICTION RESULT:**

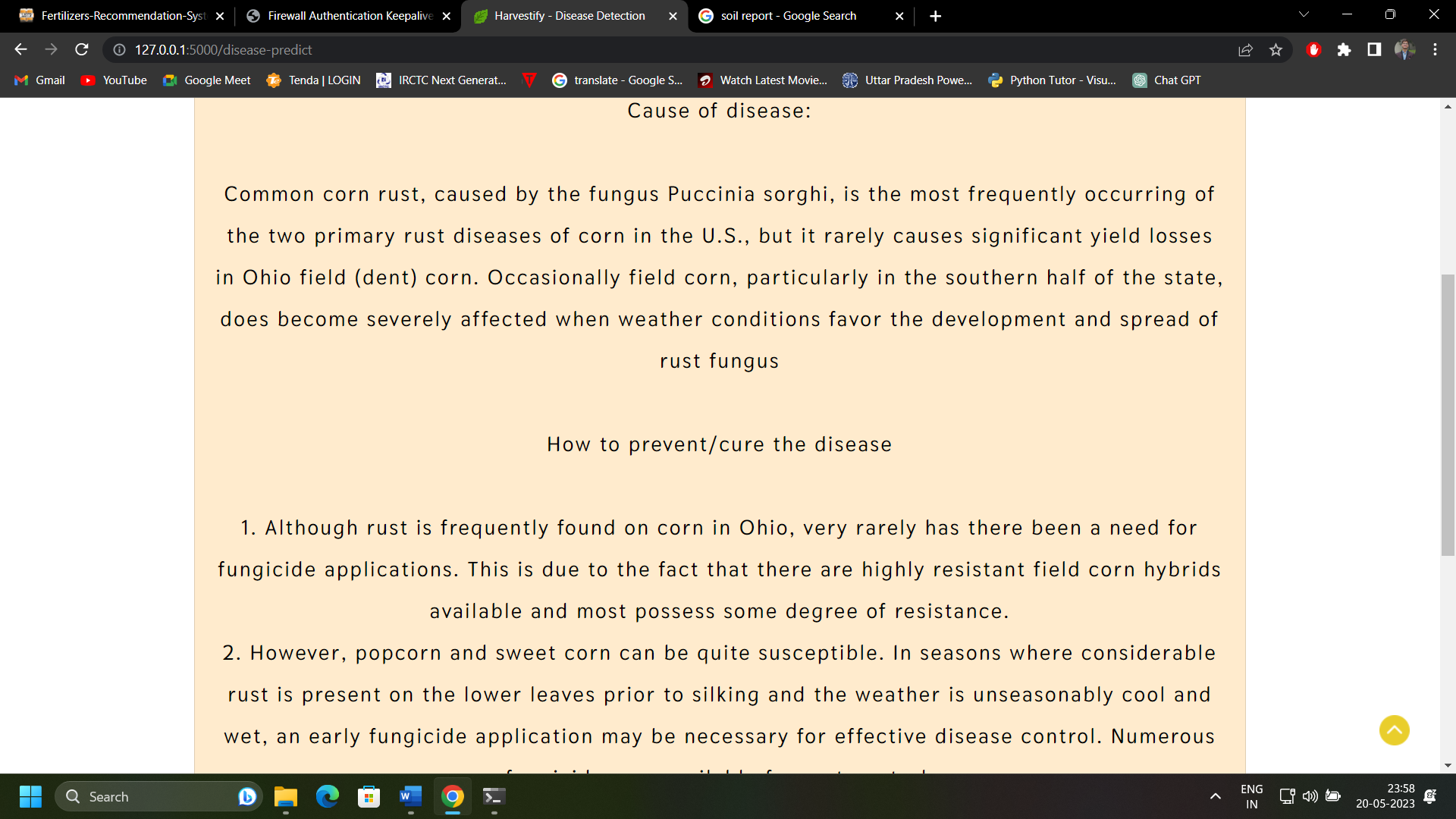
****

****

**DISEASE PREDICTION RESULT:**

****

****

****