1. INTRODUCTION

1.1 Project Overview

This project is to develop an end-to-end deep learning solution for classifying **potato leaf images into three categories**: **healthy, early blight, and late blight**. The proposed solution involves the use of convolutional neural networks to extract relevant features from the input images and classify them into one of the three categories.

1.2 Purpose

Predicting potato leaf disease at the earliest is crucial because it can have significant impacts on crop **yield, cost, and quality**. Prediction of potato leaf disease early will help the farmers to:

1. **Minimize** crop **loss**: By predicting the disease early, farmers can take timely measures to control its spread, thereby minimizing crop loss.

2. **Reduce cost**: By identifying the disease early, farmers can target the specific area of the crop affected, which helps in reducing the overall cost of control measures.

3. **Improve** crop **quality**: By predicting the disease early, farmers can take appropriate measures to prevent its spread, thereby improving the overall quality of the crop.

2. LITERATURE SURVEY

The research on agricultural development can enhance economic growth as well as can provide a healthy environment for human beings. To improve crop production, there are many deep learning model and computer vision-based studies have been conducted.

M. Islam in 2017 performed image segmentation-based potato leaf detection model and have used PlantVillage dataset. They utilized a multiclass Support Vector Machine on that segmented image to classify the diseases, finally achieved 95% accuracy.

Samajpati introduced a hybrid model to recognize the disease. The author segmented the images using k-means clustering after that classified the images using random forest algorithm. Their accuracy varied between 60 to 100%.

In 2021, Hassan Afzaal used PlantVillage dataset to classify early blight potato disease from real-time system. Here, they used few recent network model like GoogleNet, VGGNet, and EfficientNet.

Md. Khalid Rayhan Asif used CNN model to classify potato leaf images for several diseases. They divided the dataset into two classes: normal and disorder-impacted leaf. After applying five transfer learning algorithms they achieved 97% accuracy to classify the provided dataset.

For agricultural research PlantVillage dataset is one of the most popular public dataset sources. Many researches used this dataset to analyze agricultural data to improve production quality and amount.

Summary of Recent Papers for Potato Leaf Disease Prediction:

* Iqbal in 2020 used Random Forest algorithm, Private dataset of 450 images and used 2 classes of potato leaf diseases s and achieved 97% accuracy.
* Md. Khalid Rayhan Asif in 2020 used CNN algorithm, used Kaggle & Dataquest dataset, used 2 classes of potato leaf diseases and achieved 97% accuracy.
* Lee in 2021 used CNN algorithm, PlantVillage datasets, 3 classes of potato leaf diseases and achieved 99.53 accuracy.

**After observing several previous studies, I have chosen CNN algorithm, PlantVillage datasets, 3 classes of potato leaf potato leaf diseases prediction as our topic.**

***References***: Predicting and Classifying Potato Leaf Disease using K-means Segmentation Techniques and Deep Learning Networks - ([main.pdf (sciencedirectassets.com)](https://pdf.sciencedirectassets.com/280203/1-s2.0-S1877050922X0015X/1-s2.0-S1877050922016970/main.pdf?X-Amz-Security-Token=IQoJb3JpZ2luX2VjEHQaCXVzLWVhc3QtMSJHMEUCICOsvfJEhiHJK%2FxP0KU2s3Ll4wgZunUdYqMS7DAqbo%2BMAiEA1yPxMMWvVAnUCkaNUfQ5duXAxK9OUIVea8vgtOaXCSIquwUIrP%2F%2F%2F%2F%2F%2F%2F%2F%2F%2FARAFGgwwNTkwMDM1NDY4NjUiDIhTmsCNkgHvm%2B%2FV3iqPBfX25r1399U0I2PIPxYfbot%2FkxLPqqIMhtjCRG2tbD%2FaEqMV%2Ff8eUJSO1MKsvzc02McwOcOyYcehIpPzpDSusL7hzZCjuAFDve5WOTZgiEQ1sjaSMqJn00brZXvBGOgm8H6OiWKwjTGfjuUCqUT72Tj9yDLfIq8CfWDzI5Yhw35xW4S9Q%2F4akbAvFdbQZF46Vm%2BsysUUvHpPntp2lc1aeXkzQ9uPrcALQjTg%2BJTePn%2BozuCpQfVjmOtRHjNndi%2BEyW2xtdAdTi%2BZlFfG1bPbSK9my8Q4%2BrkTcmGpiUt48NkDTBCCmhuHl9VkIT236z1xo7ffy9Bt%2BuPuCQWnfc8IEtQIVejyO%2BCCI0fxRrC0pOHmZDVXZwcNsHl4bN2gP8fu0UAjhbv8Y%2FnM%2FXZ0MfhPJOTCI8wcgsLcDL%2BPI%2Baoj528v3P4ALGJXU7srfezeQQSYeQ8H5um2DrOm6bvg6XF1BW7hcb0SJ0P%2Fn4VR3DwStBgW3dPg0YWxK5S3wAFxawkqbJXuaJpcEoHxu4w2L1%2BX3AEWwDw1mmXwHTuND4tdObUlWyzm5bobckn0xX29raJ4W5S%2Bimi69S%2Bq4SMTAqbT1YnnS%2Bvo4QGPyi1QHOd9z%2FJtyVh7SabWSZgzExaxPEMPxJY8rVE2ZlcY4CBWB8qD1PP5aeNp5YDxNG26XRrOUl2YxW2ImTTPqN263E6LqWT1wb35KvAXqrt16NB0YyoNSlApflddPBXTGUAun5dGM%2FJ47mlFH5RVn1uj53h1%2FnsbwMAEUmPwPLN0QNvyDDaWLuQMytbNQzK3bLftQXmoYCY3aJujsdPgOiIAGhlOta26%2Ban3i05hETr9ZJV%2FUTxM2hCr%2Bwq97PZ9G46fBLASjUw95mqqgY6sQFpj116SS3Pu9NQs0RKTxKh4XF32RWcKr5DDsOD%2FHZKni22%2BmGORswZhRsEfLvRIIwXVLB%2FQAaDudeXUKzm%2BsZK2dcDB44q2UMWRtbZe%2Bb5jgpPToYugv7jTr%2FCWmtiMnmlJRZq5V1Yw6vbnux7Z56LhsiZheXusL2yK7dx34Q63hE7irWd5B4Hg%2Bu3zduced3Gc7r7RS4MhO9TfyGR4r0T9ianzpKPuXUeDBcqLIY8PtM%3D&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Date=20231107T194556Z&X-Amz-SignedHeaders=host&X-Amz-Expires=300&X-Amz-Credential=ASIAQ3PHCVTY277F7VIA%2F20231107%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Signature=595bbc5c1157739551c6a94d2072cd27a2d0dba1f61e6d16456fae410f2e5203&hash=c7423eefdf3c7c949444da55fec3770a617e546383974ed884aa32c5b0f97079&host=68042c943591013ac2b2430a89b270f6af2c76d8dfd086a07176afe7c76c2c61&pii=S1877050922016970&tid=spdf-69fb5100-5779-4266-9bd8-40c81defbdb7&sid=9e6c9af72ca7444ab578d1a516cc1b74e386gxrqb&type=client&tsoh=d3d3LnNja)

3. REQUIREMENT ANALYSIS

3.1 Functional requirement

The solution should classify **potato leaf images into** three categories: **healthy, early blight, and late blight**. The proposed solution involves the use of convolutional neural networks (CNNs) to extract relevant features from the input images and classify them into one of the three categories and the solution should have:

1. **Accurate prediction**: The predictor must be able to accurately predict the stage of leaf degradation. The accuracy of the prediction is crucial for farmers, agribusinesses, and other stakeholders to make informed decisions on the production.

2. **User-friendly interface**: The predictor must have a user-friendly interface that is easy to navigate and understand. The interface should present the results of the predictor in a clear and concise manner to enable farmers and other stakeholders to make informed decisions.

3.2 Non-Functional requirements

**Scalability**: The solution must be able to scale up based on the prediction from our product. The model should be able to handle any size of data without compromising on its accuracy or efficiency.

4. PROJECT CONFIGURATION STRUCTURE

Project folder as given below to store various artifacts related to the project with respective details:

***POTATOLEAFDESEASEDETECTION***

* ***> Models (***Model folder contains saved models for the project.)
* ***> Training (***Training folder contains project data and code for building/training/testing the model)
* ***> Web.py (***Web.py folder for Streamlit application python script for a website)

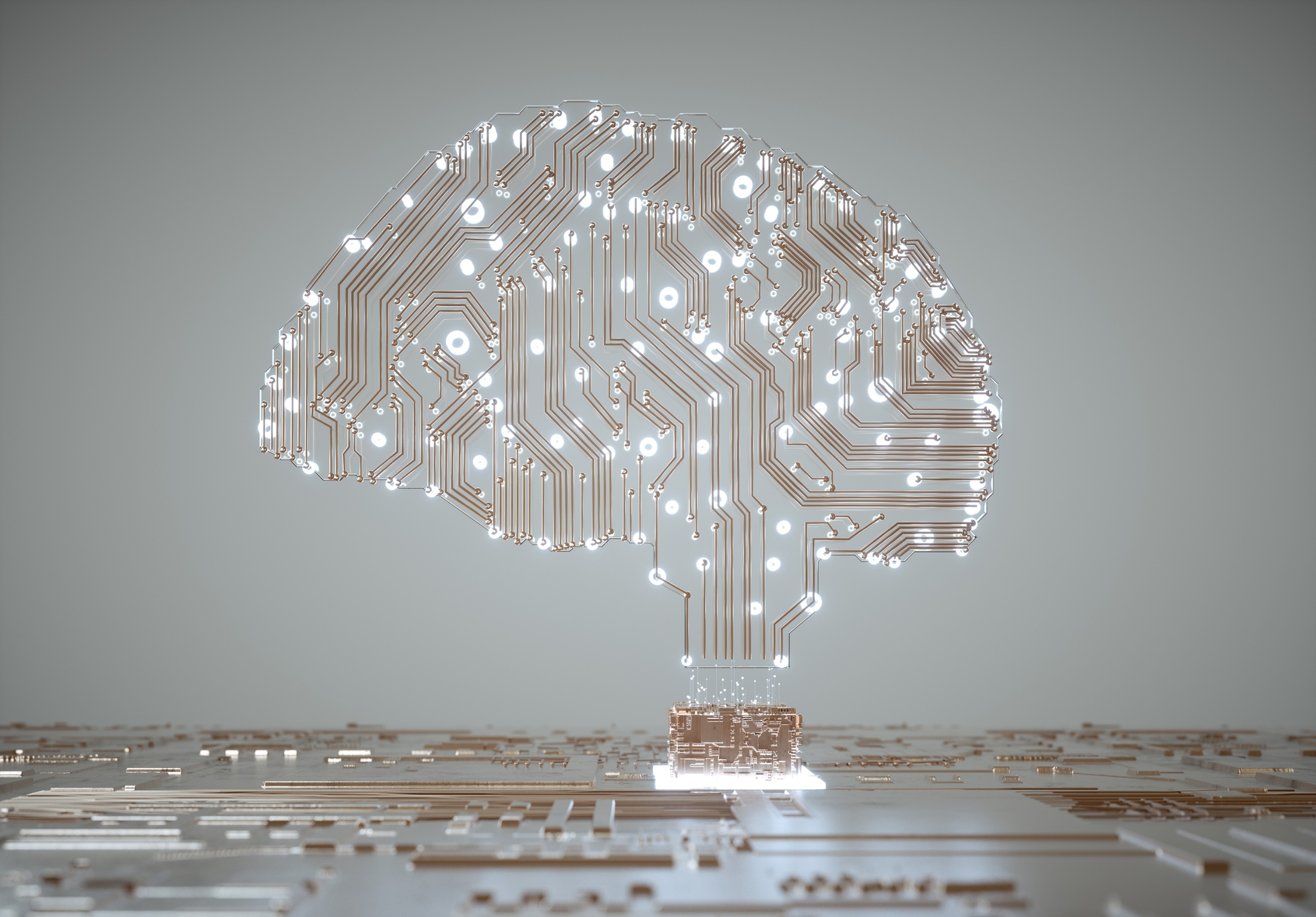
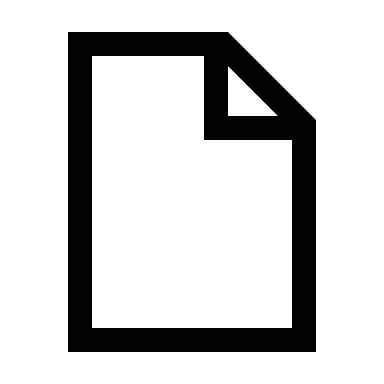
5. SOLUTION DESIGN

5.1. Technical Architecture



**Prediction**

**Train Data**



**Data**

**UI**

**Data Pre-processing**

**Model**

**Algorithm**





**Inputs**

**Test Data**

**User**

6. CODING & SOLUTIONING (Explain the features added in the project along with code)

6.1 Feature 1

6.2 Feature 2

7. PERFORMANCE TESTING

7.1 Performance Metrics

8. RESULTS

8.1 Output Screenshots

9. CONCLUSION

This project has used a deep learning-based method to classify potato leaf diseases. For performing this project, used the dataset from PlantVillage. Thereafter, used several pre-processing steps to achieve a desired outcome. We used Convolutional Neural Network concepts processing the data. We have also created a web application using Streamlit framework for the end user.

The project solution can help the farmers optimize their potato crop yields by providing accurate results. This can lead to increased crop yields and better crop quality, which can positively impact the income of farmers and the availability of good quality potatoes for consumers. The solution can also help farmers reduce production costs by providing insights into optimal planting. This can lead to better resource allocation, reduced waste, and increased profitability for farmers. Also help to reduce the usage pesticides which is good for the consumers and environment.

10. FUTURE SCOPE

In the future work, we will create a tool to identify the type of leaf disease and incorporate other algorithms to improve the performance of the model. This will help farmers working in the agricultural sector to recognize particular diseases early on, which will enable them to take the appropriate action. By implementing certain new algorithms, we will enhance the research and to improve the results, utilize a wide range of dataset types and classifications. Lastly, provide a more richer application for crop fields.

***Source Code GitHub & Project Demo link***