Project Title- Automated Potato Disease Recognition

**Abstract**

**Project Objectives-**This Project aims to the classification of potato diseases using a Convolutional Neural Network (CNN). The goal is to identify and categorize potato leaf diseases to assist farmers in making informed decisions for disease prevention and crop protection. A dataset comprising 10,000 images of potato leaves was collected from Kaggle (PlantVillage). The dataset includes three classes: Potato Early Blight, Potato Late Blight, and Healthy Potato Leaves. The research involves key steps such as data acquisition, preprocessing, augmentation, and CNN-based image classification.

**Methodology**

1. Data Collection – A dataset of potato leaf images was obtained from Kaggle, the dataset contains three disease categories: Early Blight, Late Blight, and Healthy Leaves.
2. Data Preprocessing – The collected images were resized, normalized, and converted into a format suitable for CNN training.
3. Data Augmentation – Techniques like rotation, flipping, zooming, and brightness adjustment were applied to enhance the dataset’s variability and improve model generalization.
4. Model Training – A CNN model was trained on the dataset using different epoch settings (30, 40, and 50 epochs) to evaluate performance.
5. Evaluation Metrics – The model was evaluated based on accuracy, loss, precision, recall, and F1-score to determine its effectiveness in classifying potato diseases.

**Key findings**

1. The CNN model achieved 100% accuracy at 40 epochs, proving its efficiency in classifying potato diseases.
2. The model also demonstrated 99.97% accuracy at 30 epochs and 99.98% accuracy at 50 epochs, but 40 epochs provided the best balance between training time and performance.
3. Data augmentation significantly enhanced the model’s ability to generalize across various image conditions.
4. The study confirms that CNN-based classification can play a vital role in early disease detection, aiding farmers in timely intervention and crop protection.

Step wise solution approach-

Step 1: Data Collection

Step 2: Data Preprocessing and Cleaning (tf dataset and Data Augmentation)

Step 3: Model Building (CNN)

Step 4: Model Training

Step 5: Model Evaluation

Step 6: Model Optimization and Hyperparameter Tuning

Reference:

[1] https://hal.science/hal-04015255/document

Team:

1. Shivam Shukla (E22CSEU1443)