**National University of Computer & Emerging Sciences**

**Karachi Campus**



Potato Disease Classification

**Deep Learning for Perception**

**Section: A**

**Group Members:**

**21K-3262 Alyan Khowaja**

**21k-3326 Syed Hadi Arshad**

**21k-**

1. Abstract:

This project focuses on the development of a system for the classification of diseases in potatoes using machine learning. The goal is to detect and diagnose different diseases that affect potatoes thus preventing severe consequences. For this purpose, a dataset with images of healthy and infected potato plants was collected. Furthermore, the data was preprocessed and used to train different deep learning models, allowing for their classification. The outcomes indicate the efficiency of the approach that reached a 95% accuracy on the test set.

2. Introduction:

The potato disease classification project aims to use the capabilities of CNNs for the early detection and classification of diseases that plague potato plants. The whole world is depending on potato as a basic food, so, the health and productivity of potato is very important for the food security. Nevertheless, there are numerous diseases, for instance, the early and the late blight, that are major threats to the potato production if they are not detected and treated in time. The project is set to solve this problem by creating an automated classification methodology that can classify the images of potato leaves and correctly identify the symptoms of the disease.

3. Problem Statement:

The main problem that this project is going to solve is the fast detection of the diseases which affects potato plants. The first thing that comes to my mind when seeing diseases like early blight and late blight is that of identification at the early stage which is very important for the implementation of interventions aimed at the reduction of their effects and also to decrease the yield losses. Artsically, the conventional ways of identifying the diseases are the visual inspection by the experts which is a long and subjective process. Hence, there is a definite necessity for the automated systems that can perfectly detect the disease symptoms by analyzing the images.

4. Objectives:

* Create a CNN-based classification model that will be able to correctly classify the potato diseases based on the images of plant leaves.
* Within the scope of this project, the evaluation of the developed model will be done on a diverse dataset having both healthy potato leaves and leaves affected by early blight and late blight.
* The collected model can be utilized for agriculture, where it will be applied for the early detection of diseases, following the progression of the disease and providing the targeted treatment strategies.

5. Methodology:

The methodology employed in this project involves several key steps:

* Data Collection: The dataset is from Kaggle, which is a collection of pictures of healthy potato leaves, early blight leaves, and late blight leaves.
* Data Preprocessing: Refurbishing the dataset and augmenting its quality through the usage of TensorFlow datasets to boost the model performance and adjustability. The techniques which are under consideration as part of the resizing, rescaling, and data augmentation include, the diversification of the training samples.
* Model Development: The design and implementation of the CNN architecture for potato disease classification is the aim of this research. This is framed in the process of determining the layers, activation functions, loss functions, optimizers and backpropagation mechanisms of the architecture.
* Training and Evaluation: The process of training the CNN model on the preprocessed dataset and assessing its performance with the help of the metrics like accuracy, sensitivity, false positive and false negative. Besides, the relation with the alternative algorithms like Random Forest classifier and K Nearest Neighborhood for the comparison of the results.

6. System Development:

The process of writing the details of the CNN model architecture, including the layout of the layers, the activation functions, and the optimization techniques that have been used are included. Data preprocessing, model building, optimization, and validation are the main steps of the system development process, which is in charge of the high reliability and robustness of the classification method that is developed.

7. Performance Analysis:

The critical examination of the model's performance is done through the evaluation of the accuracy metrics, sensitivity analysis, and the false positive and false negative rates. The comparative analysis of the CNN-based approach is a proof of the effectiveness of the technique in the potato disease classification. The, through performance analysis, a broader view of the project is gained, and eventually, the project and the feedback from it are concluded and improved.

8. Conclusions:

A briefing of the main achievements and the consequences of the results obtained through the project is the description of the summary of the key findings and implications that are derived from the project's outcomes. The conclusions underline the importance of the early disease detection in potato plants, the effectiveness of the CNN-based classification systems and also, the possible future research and development in the field of agricultural automation.

9. Future Scope:

The next steps the project will take should also be discussed, such as how to make the model more accurate, scalable and how to deploy it in the real world. In future, this also includes the creation of various auxiliary tools and mobile apps that allow on-site disease monitoring and public access to the disease classification services through cloud-based platforms.