## Plant Disease Prediction Model Report

**Introduction**

Gardening enthusiasts and plant lovers often encounter challenges in identifying plant diseases accurately, leading to frustration and potential plant loss. Similarly, businesses like plant retailers and nurseries face difficulties in maintaining plant health and minimizing losses due to disease outbreaks. To address these issues, a plant disease prediction model was developed and deployed using a Convolutional Neural Network (CNN) trained with 38 classes of plant diseases. This report provides an overview of the model, its performance metrics, deployment details, challenges faced, and a demonstration of the application.

**ML Canvas**



**Model Benchmarking**

The CNN model was trained with 38 classes, employing increasing filters, dropout layers, and dense layers, along with the Adam optimizer with a learning rate of 0.0001. Evaluation metrics such as accuracy, precision, recall, and F1-score were computed for each class, achieving a training accuracy of around 95.82% and a validation accuracy of around 92.59%. These metrics highlight the model's ability to classify plant diseases accurately across a diverse range of classes.

A screenshot of a computer screen

Description automatically generated

* Accuracy and Confusion matrix analysis

A graph of a graph showing the performance of a performance

Description automatically generated with medium confidence A screenshot of a computer screen

Description automatically generated

**Deploying Plant Disease Prediction Model**

The model was deployed using the Streamlit framework (v1.30.0) for rapid development and deployment. Compatibility with required versions of NumPy (v1.26.3) and TensorFlow (v2.15.0.post1) was ensured for seamless integration. Local testing and validation were conducted thoroughly to validate functionality and responsiveness. The application was successfully deployed on the Streamlit Sharing platform, allowing real-time predictions of plant diseases.

**Challenges & Issues Encountered**

Several challenges were encountered during the development and deployment of the plant disease prediction model:

1. Data Imbalance: Dealing with imbalanced classes in the dataset required strategies to mitigate biased model predictions and maintain overall performance metrics.

2. Overfitting: Balancing model complexity and generalization was crucial to avoid overfitting. A small learning rate (0.0001) was chosen to prevent overshooting.

3. Hyperparameter Tuning: Experimenting with hyperparameters such as learning rates, batch sizes, number of epochs, and dropout rates was time-consuming and required significant computational resources.

**Conclusion**

The plant disease prediction model addresses a critical need for both individual customers and businesses in accurately identifying and managing plant diseases. With high accuracy and robust performance metrics, the model provides valuable insights for early disease detection, informed purchasing decisions, and educational purposes. Despite challenges such as data imbalance, overfitting, and hyperparameter tuning, the successful deployment of the model on the Streamlit platform ensures accessibility and usability for users seeking to enhance plant health and mitigate economic losses.

**Reference**

Dataset: [New Plant Diseases Dataset (kaggle.com)](https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset/data)