

Data Collection and Preprocessing Phase

Date	15 March 2024
Team ID	SWTID1720033149
Project Title	Visual Diagnostics: Detecting Tomato Plant Diseases With Leaf Image Analysis
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan Template

Section	Description
Project Overview	The "Visual Diagnostics: Detecting Tomato Plant Diseases Through Leaf Image Analysis" project aims to develop an intuitive diagnostic tool that uses advanced machine learning to identify tomato plant diseases from leaf images. This tool, accessible via a user-friendly mobile application, will enable farmers to quickly and accurately diagnose diseases, providing instant results and actionable treatment recommendations. By leveraging deep learning techniques and offering a reliable, cost-effective solution, the project seeks to improve crop health, increase yields, and enhance farmers' satisfaction by reducing reliance on subjective visual inspections and costly lab tests, ultimately promoting sustainable farming practices.

Data Collection Plan	<p>The dataset that has been used for this project has been collected from Kaggle(https://www.kaggle.com/datasets/kaustubhb999/tomatoleaf)</p> <p>The images of tomato leaves have been collected and organized into subdirectories based on their respective names as per the project structure.</p>
Raw Data Sources Identified	<p>Tomato Leaf Image Dataset (Kaggle): Contains labelled images of tomato leaves in 10 categories, used for training and validating the model (Kaggle - Tomato Leaf Image Dataset).</p> <p>GitHub Link: Backup of the project notebook, accessible via a link in the project's GitHub repository.</p>

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Tomato Leaf Image Dataset (Kaggle)	Contains labelled images of tomato leaves in 10 categories, used for training and validating the model.	https://www.kaggle.com/datasets/kaustubhb999/tomatoleaf	Image	188 MB	Public

Google Drive	Backup of the project notebook and resources for sharing.	https://drive.google.com/drive/folders/1qnGJNgSl_ZvwmwsLgbac7IeSw6BNLwvQ?usp=sharing	Folder		Anyone With the Link may View
...