

Data Collection and Preprocessing Phase

Date	20 July 2024
Team ID	SWTID1720109344
Project Title	Rice Type Classification Using Cnn
Maximum Marks	2 Marks

Data Quality Report Template

The Data Quality Report Template will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Data Source	Data Quality Issue	Severity	Resolution Plan
Dataset	Mention the issues faced in the selected dataset.	Low/ Moderate / High	Give the solution for that issue technically.
Kaggle Rice Image Dataset	Inconsistent Image Sizes	Moderate	Implement preprocessing to resize all images to a uniform size (e.g., 224x224 pixels) using libraries like OpenCV or PIL.
Kaggle Rice Image Dataset	Image Noise and Artifacts	Low	Apply denoising filters, such as 'cv2.fastNlMeansDenoisingColor', to clean the images and improve clarity.

Kaggle Rice Image Dataset	Class Imbalance	Low	Use data augmentation (e.g., rotation, flipping) to balance the number of samples per class. Apply class weighting or oversampling techniques during model training.
---------------------------------	-----------------	-----	---