**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 12 July 2024 |
| Team ID | SWTID1720157891 |
| 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 | Insufficient Diversity of Rice Grain Types | Moderate | Collected additional images representing a broader range of rice grain types. Augmented existing dataset with synthetic data or data from different sources to increase diversity. Ensured balanced representation across all classes to improve model generalization. |
| Dataset | Image quality Variations | High | Implemented image preprocessing techniques such as normalization, resizing, and noise reduction to standardize image quality. Applied data augmentation to simulate variations in lighting, angle, and quality to make the model robust to such variations during training. |