**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 20 June 2025 |
| Team ID | SWTID1749826875 |
| Project Title | Dog Breed Identification using Transfer Learning |
| Maximum Marks | 2 Marks |

**Data Quality Report**

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** |
| Kaggle Dataset 1 | The testing data was unlabelled. | Moderate | Separate data were used to verify the accuracy of the model |
| Kaggle Dataset 1 | The dataset had more than required classes | Low | The data of required classes were filtered |
| Kaggle Dataset 1 | The data were not classified into folders | Low | The data were filtered into folders using os module functions |
| Kaggle Dataset 2 | The labels were not matching with training dataset | Low | Used string modification functions to normalize the labels |