**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 12 July 2024 |
| Team ID | SWTID1719935963 |
| Project Title | Automated Weather Classification using Transfer Learning |
| 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** |
| Weather Image Classification Dataset | Limited Data Size (1500 images) | Moderate | Explore data augmentation techniques like random flips, rotations, or color jittering to artificially increase dataset size. Consider using transfer learning from pre-trained models on larger image datasets. |
| Weather Image Classification Dataset | Missing Train/Test Split | High | Split the data into training and testing sets using a common approach like random sampling. A typical split might be 80% for training and 20% for testing. Stratified sampling can be considered if there's class imbalance. |