**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 15 March 2024 |
| Team ID | SWTID1720707508 |
| Project Title | WarLens: Transfer Learning for Event Classification in Conflict Zones |
| 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** |
| https://www.kaggle.com/datasets/saailna/war-events-classification | Class Imbalance | High | Apply techniques like oversampling, undersampling, or using SMOTE to balance class distributions. |
|  | Noise in Data | Moderate | Manually review images to remove irrelevant ones or use automated filters to clean noisy data. |
|  | Insufficient Metadata | Moderate | Enrich the dataset with additional metadata such as event context, location, and date |