 

Data Collection and Preprocessing Phase

|  |  |
| --- | --- |
| Date | 03 October 2024 |
| Team ID | LTVIP2024TMID24818 |
| Project Title | Toxic Comments Classification For Social Media |
| Maximum Marks | 2 Marks |

**Data Quality Report:**

The Data Quality Report 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 Quality Report:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Source** | **Data Quality Issue** | **Severity** | **Resolution Plan** |
| Kaggle Dataset | Categorical data in the dataset | Moderate | encoding has to be done in the data. |