 

Data Collection and Preprocessing Phase

|  |  |
| --- | --- |
| Date | 15 July 2024 |
| Team ID | 739716 |
| Project Title | Predicting Baseline Histological staging in hcv patients using machine 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 | No null values in the dataset | Moderate | - |

|  |  |  |  |
| --- | --- | --- | --- |
| Kaggle dataset | Catageorical data in the dataset | moderate | Encoding has to be done in the data |