**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 15 March 2024 |
| Team ID | SWTID1749727925 |
| Project Title | E-Adapt: Predicting Student Adaptability in Online Classes |
| Maximum Marks | 2 Marks |

**Data Quality Report Template**

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Source** | **Data Quality Issue** | **Severity** | **Resolution Plan** |
| Students Adaptability Level in Online Education (Kaggle) | Missing values in some features | Low | |  | | --- | |  |  |  | | --- | | Impute missing values with mode or drop rows if minimal. | |
| Students Adaptability Level in Online Education (Kaggle) | Duplicate rows found | Low | Remove duplicates using drop\_duplicates() in Pandas. |
| Students Adaptability Level in Online Education (Kaggle) | Imbalanced target classes(if any) | Moderate | Use class weighting or SMOTE for balancing during model training. |
| Students Adaptability Level in Online Education (Kaggle) | Inconsistent category labels (e.g., lowercase/uppercase) | Low | Standardize category labels using string methods. |