 

Data Collection and Preprocessing Phase

|  |  |
| --- | --- |
| Date | 04TH July 2024 |
| Project Title | AI Feedback Generator For Development Teams Using Palm's Text-Bison-001 |

**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** |
| Public Repositories (GitHub) | Incomplete code snippets or projects | High | Filter out incomplete submissions or request additional information from the repository owners. |
| Public Repositories (GitHub) | Inconsistent coding styles across different submissions | Moderate | Implement a standardization process for coding styles using a linter before analysis. |
| Public Repositories (GitHub) | Missing metadata about the projects (e.g., project descriptions, tags) | Moderate | Use web scraping and API integration to fetch additional metadata from the repositories. |