**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 15 March 2024 |
| Team ID | INDIVIDUAL |
| Project Title | CareerCraft: ATS-Optimized Resume Analyzer using Gemini Model |
| 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** |
| CareerCraft Project Database | Incomplete Job Descriptions | High | Implement a pre-processing step to handle incomplete data by filling in missing fields using industry-standard defaults or by flagging them for manual review. |
| Job description | eyword Matching | Moderate | Standardize the text processing pipeline to ensure consistent keyword extraction and matching, including the use of NLP techniques to handle synonyms and related terms. |