Data Collection and Preprocessing Phase

Date	15 July 2024
Team ID	739718
Project Title	Polycystic Ovary Syndrome Classification Using Machine Learning
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
Polycystic Ovary Syndrome Dataset	Incomplete Data	Moderate	Use data imputation methods to fill missing values.
Polycystic Ovary Syndrome Dataset	Duplicate Entries	Moderate	Identify and remove duplicate records

Polycystic Ovary Syndrome Dataset	Outdated Data	Low	Supplement with recent data or apply temporal analysis.
Polycystic Ovary Syndrome Dataset	Erroneous Data	High	Validate data entries and correct errors using predefined rules.
Polycystic Ovary Syndrome Dataset	Inconsistent Data	High	Standardize naming conventions and clean the data.