

Data Collection and Preprocessing Phase

Date	23 September 2024
Team ID	LTVIP2024TMID24992
Project Title	Rainfall Prediction 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
Smart Internz Dataset	Missing values in the "MinTemp","MaxTemp","Rainfall","Evaporation","Sunshine","WindGustDir","WindGustSpeed","WindDir9am","WindDir3pm","WindSpeed9am","WindSpeed3pm","Humidity9am","Humidity3pm","Pressure9am","Pressure3pm","Cloud9am","Cloud3pm","Temp9am","Temp3pm","RainToday","RainTomorrow"	Moderate	Use mean/mode Imputation

Smart Internz Dataset	Categorical data in the dataset	Moderate	Encoding has to be done in the data
-----------------------------	---------------------------------	----------	--