

Data Collection and Preprocessing Phase

Date	04 November 2024
Team ID	SWTID1726834817
Project Title	Fake News Analysis in social media using NLP
Maximum Marks	2 Marks

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
Kaggle Dataset	Missing values in the "text" column.	Low	Add code to check for missing values using <code>df['text'].isnull().sum ()</code>
Kaggle Dataset	Class Imbalance	Moderate	Use <code>df['label'].value_counts()</code>