Data Collection and Preprocessing Phase

Date	June 2024
Team ID	740107
Project Title	The Language Of You tube: A Text Classification Approach To Video Description
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: Arduino Database

Data Source	Data Quality Issue	Severity	Resolution Plan
YouTube Data API	Variations in text formatting (e.g., case sensitivity, different spellings, abbreviations)	High	Use the YouTube Data API to fetch comments from a variety of videos