

Data Collection and Preprocessing Phase

Date	15 March 2024
Team ID	SWTID1720333657
Project Title	Wce Curated Colon Disease Classification Using Deep
Maximum Marks	2 Marks

Data Quality Report Template

The data originated from a Kaggle dataset, which generally implies a level of usability. However, we identified opportunities to further enhance the data quality. We employed a technique called data augmentation, which essentially creates additional data points from existing ones. This approach effectively bolsters the dataset's robustness without requiring substantial modifications to the original information.

Data Source	Data Quality Issue	Severity	Resolution Plan
Kaggle dataset	No issues faced. Just had to ameliorate the quality of raw data.	Low	With help of data augmentation.
...