

## Data Collection and Preprocessing Phase

Date	15 October 2024
Team ID	LTVIP2024TMID24772
Project Title	Implementation of Deep Learning Techniques to Detect Malaria
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
Kaggle Dataset	<b>Label Encoding Mismatch:</b> Labels for infected and uninfected classes may not be properly encoded.	Moderate	Apply <b>one-hot encoding</b> to transform the categorical labels into a binary matrix. This ensures the model interprets the labels correctly for multi-class classification.
Kaggle Dataset	<b>Inconsistent Pixel Value Range:</b> Pixel values in images	Low	Use <b>normalization</b> to scale the pixel values between <b>0 and 1</b> by dividing them by 255. This improves the training process and

	may vary (e.g., range 0-255) causing issues.		ensures faster model convergence.
--	---	--	--------------------------------------