Data Collection and Preprocessing Phase

Date	25 JUNE 2025
Team ID	SWTID1749974387
Project Title	**Neural Networks Ahoy: Cutting-Edge Ship Classification For Maritime Mastery**
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	Class imbalance and limited variety in lighting/angles/background	Medium	Used ImageDataGenerator with rescaling, zoom, shear, and horizontal flipping to augment data and improve model generalization. Also applied an 80-20 training-validation split to monitor overfitting.
Kaggle Dataset	Categorical data in the dataset	Moderate	Mapping has to be done in the data.