

## Data Collection and Preprocessing Phase

Date	15 June 2025
Team ID	SWTID1750006853
Project Title	ASL- Alphabet Image Recognition
Maximum Marks	2 Marks

### Data Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

### Data Collection Plan Template

Section	Description
Project Overview	The ASL Alphabet Image Recognition project aims to build a deep learning model capable of recognizing static hand gestures representing American Sign Language alphabets. This will facilitate easier communication and serve as a tool for ASL learners and the hearing-impaired community.
Data Collection Plan	The dataset for this project is sourced from Kaggle, a public data platform. It includes labeled images for each ASL alphabet letter and some special gestures. The data is already pre-collected and annotated, so no manual labeling or additional scraping is required.
Raw Data Sources Identified	The primary source is the "ASL Alphabet" dataset, which contains over 87,000 images sorted into 29 classes (A–Z, space, delete, nothing). It is well-structured and publicly accessible via Kaggle.

### Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
ASL Alphabet Dataset	Contains labeled images of ASL hand gestures for 26 alphabets and 3 additional signs (space, nothing, delete). Images are organized in subfolders per class.	<a href="https://www.kaggle.com/datasets/grassknoted/asl-alphabet">https://www.kaggle.com/datasets/grassknoted/asl-alphabet</a>	JPEG (Images)	~2 GB	Public (Kaggle account required)