

## Data Collection and Preprocessing Phase

Date	10 December 2024
Team ID	739902
Project Title	Alzheimer Disease Prediction
Maximum Marks	6 Marks

### Preprocessing Template

The images will be preprocessed by resizing, normalizing, augmenting. These steps will enhance data quality, promote model generalization, and improve convergence during neural network training, ensuring robust and efficient performance across various computer vision tasks.

Section	Description
Data Overview	<ol style="list-style-type: none"><li>1. The dataset includes brain MRI images.</li><li>2. It contains many records, with images in JPG format.</li></ol>
Resizing	Resize brain MRI images to a standard resolution to ensure uniform input size.
Normalization	Normalize pixel values of MRI images to a range between 0 and 1 for consistency.
Data Augmentation	Apply augmentation techniques like random flipping, rotation, and zooming to increase data variability and prevent overfitting.

## Data Preprocessing Code Screenshots

### Loading Data



```

jupyter Alzheimer Last Checkpoint: 23 hours ago
File Edit View Run Kernel Settings Help
JupyterLab Python 3 (ipykernel)

[ ]: ##DataCollection
[1]: import zipfile
import os

# Path to the uploaded ZIP file
zip_file_path = 'archive.zip'

# Folder where you want to extract the ZIP file
extract_folder = 'extract_files'

# Create the folder if it doesn't exist
os.makedirs(extract_folder, exist_ok=True)

# Extract the contents of the ZIP file
with zipfile.ZipFile(zip_file_path, 'r') as zip_ref:
    zip_ref.extractall(extract_folder)

print(f'Files extracted to: {extract_folder}')

Files extracted to: extract_files

```

### Resizing, Normalization and Data Augmentation



```

[ ]: ##Configure of Images & Preprocessing

[7]: from tensorflow.keras.preprocessing.image import ImageDataGenerator as IDG
IMG_SIZE = 180
IMAGE_SIZE = [180, 180]
DIN = (IMG_SIZE, IMG_SIZE)
ZOOM = [.99, 1.01]
BRIGHT_RANGE = [0.8, 1.2]
HORIZ_FLIP = True
FILL_MODE = "constant"
DATA_FORMAT = "channels_last"
WORK_DIR = "extract_files/Alzheimer_s Dataset/train"
work_dir = IDG(rescale = 1./255, brightness_range = BRIGHT_RANGE, zoom_range = ZOOM, fill_mode = FILL_MODE, data_format = DATA_FORMAT, horizontal_flip =
train_data_gen = work_dir.flow_from_directory(directory = WORK_DIR, target_size = DIN, batch_size = 6500, shuffle = False)

Found 5121 images belonging to 4 classes.

```