

Data Collection and Preprocessing Phase

Date	15 March 2024
Team ID	Team-738168
Project Title	Cognitive Care: Early Intervention for Alzheimer's Disease
Maximum Marks	6 Marks

Preprocessing Template

Preprocessing is a crucial step in preparing your data for analysis when investigating the effectiveness of cognitive care programs for early-stage Alzheimer's Disease (AD). Analyze the data to identify the extent and patterns of missing values across different variables. Consider the reasons for missing data. If missingness is related to the variable itself (e.g., patients with severe cognitive decline might not complete cognitive tests), imputation techniques may not be ideal. Explore alternative approaches like excluding participants with excessive missing data. Identify and rectify errors in the data, such as typos, outliers, or inconsistencies in date formats or units. Utilize data validation techniques to catch potential issues.

Section	Description
Data Overview	The dataset for the project consist of images which contain Alzheimer's Diseases MRI Images. Also contain 4 stages of Alzheimer's Diseases i.e MildDemented, VeryMildDemented, NonDemented, ModerateDe mented.
Resizing	Resizing the images to the target size i.e target_size = (224,224)
Normalization	Normalize pixel using preprocessing_input function to convert image from RGB to BGR, as the standard format of imagenet

Data Augmentation	rotation_range, horizontal_flip, width_shift_range, height_shift_range, validation_split and preprocessing_function.
Data Preprocessing Code Screenshots	
Loading Data	<pre>[1] !unzip '/content/archive (11).zip'</pre> <pre> inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem947.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem948.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem949.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem95.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem950.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem951.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem952.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem953.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem954.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem955.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem956.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem957.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem958.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem959.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem96.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem960.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem961.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem962.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem963.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem964.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem965.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem966.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem967.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem968.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem969.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem97.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem970.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem971.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem972.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem973.jpg inflating: Alzheimer_s Dataset/train/VeryMildDemented/verymildDem974.jpg </pre> <pre>[3] train_data_dir = '/content/Alzheimer_s Dataset/train' test_data_dir = '/content/Alzheimer_s Dataset/test'</pre>
Resizing	<pre> ▶ train_dataset = train_datagen.flow_from_directory(directory = '/content/Alzheimer_s Dataset/train', target_size = (224,224), class_mode = 'categorical', subset = 'training', batch_size = 128) Found 4098 images belonging to 4 classes. ▶ valid_dataset = valid_datagen.flow_from_directory(directory = '/content/Alzheimer_s Dataset/train', target_size = (224,224), class_mode = 'categorical', subset = 'validation', batch_size = 128) Found 1023 images belonging to 4 classes. </pre>

Normalization

```
train_datagen = ImageDataGenerator(rescale = 1./255,  
                                   rotation_range=30,  
                                   zoom_range=0.2,  
                                   horizontal_flip=True,  
                                   vertical_flip=True,  
                                   validation_split = 0.2)  
  
valid_datagen = ImageDataGenerator(rescale = 1./255,  
                                   validation_split = 0.2)  
  
test_datagen = ImageDataGenerator(rescale = 1./255)
```

Data Augmentation

```
train_datagen = ImageDataGenerator(rescale = 1./255,  
                                   rotation_range=30,  
                                   zoom_range=0.2,  
                                   horizontal_flip=True,  
                                   vertical_flip=True,  
                                   validation_split = 0.2)  
  
valid_datagen = ImageDataGenerator(rescale = 1./255,  
                                   validation_split = 0.2)  
  
test_datagen = ImageDataGenerator(rescale = 1./255)
```