

Data Collection and Preprocessing Phase

Date	20 October 2024
Team ID	739652
Project Title	Ai-Powered Nutrition Analyzer For Fitness Enthusiasts
Maximum Marks	6 Marks

Preprocessing Template

Preprocessing involves preparing data for analysis by cleaning, transforming, and organizing it. Steps include handling missing values, normalizing data, encoding categorical variables, and ensuring consistency and accuracy. It enhances the quality and usability of data for machine learning and other analytical tasks.

Section	Description
Data Overview	Provides a summary of key nutritional data, including macronutrients, micronutrients, and dietary preferences.
Resizing	Adjusts input data dimensions to fit specific fitness goals, such as meal plans, portion sizes, or dietary trends.
Normalization	Standardizes data by converting all nutritional values into a common scale, ensuring consistent comparisons.

Data Augmentation	Enhances nutritional data through techniques like generating variations of meal plans using healthy substitutions.
Denoising	Removes noise from nutritional data, such as irrelevant ingredients or incorrect quantities, for improved accuracy.
Edge Detection	Extracts key nutrients or combinations, like protein-carbohydrate ratios or vitamins-minerals combinations.
Color Space Conversion	Converts ingredient data into embeddings (e.g., Nutrient Embeddings) for richer analysis.
Image Cropping	Trims meal data to include only necessary components, such as primary ingredients or essential macros.
Batch Normalization	Normalizes nutritional data frequencies for balanced meal proportions, e.g., protein, carbs, fats percentages.
Data Preprocessing Code Screenshots	
Loading Data	<pre>x_train = train_datagen.flow_from_directory(r"/content/drive/MyDrive/TRAIN_SET/TRAIN_SET", target_size=(64, 64), batch_size=5, color_mode='rgb', class_mode='sparse') x_test = test_datagen.flow_from_directory(r"/content/drive/MyDrive/TRAIN_SET/TEST_SET", target_size=(64, 64), batch_size=5, color_mode='rgb', class_mode='sparse')</pre>

Coreleation between variables

```
img=image.load_img(r"/content/drive/MyDrive/TRAIN_SET/TEST_SET/BANANA/103_100.jpg",
                    color_mode='rgb',target_size=(64,64))
x=image.img_to_array(img)
x=np.expand_dims(x,axis=0)
pred=np.argmax(model.predict(x),axis=-1)
pred
```

Data Preprocessing	<pre> classifier = Sequential() classifier.add(Conv2D (32, (3, 3), input_shape=(64, 64, 3), activation='relu')) classifier.add(MaxPooling2D(pool_size=(2, 2))) classifier.add(Conv2D (32, (3, 3), activation='relu')) classifier.add(MaxPooling2D(pool_size=(2, 2))) classifier.add(Flatten()) </pre>
clean the comment _text in both the datasets. & training and testing	<pre> classifier.add(Dense (units=128, activation='relu')) classifier.add(Dense (units=5, activation='softmax')) </pre>
Train _test_ split	<pre> train_datagen=ImageDataGenerator(rescale=1./255,shear_range=0.2,zoom_range=0.2,horizontal_flip=True) test_datagen=ImageDataGenerator(rescale=1./255) </pre>
Loading the pickle file	<pre> classifier.save('nutrition.h5') </pre>