A screenshot of a computer program

Description automatically generated

**NOTE: The above screenshot tells that we have scraped recipes from different links randomly from website tarladal.com**

**We used the method to get the last column from excel to write the recipes from where we stopped during last run. So, it will not overwrite.**

**Just changing different pages link as above screenshot we can be able to scrape more recipes satisfying add, eliminate, allergies/Nut allergies filters**

**We assumed that given allergy ingredients should be matched for all the conditions. (Allergy recipes excel has recipes with allergy matched ingredients for Diabetes/Hypothyroidism/Hypertension/PCOS)**

**We make sure the recipes do not have eliminated ingredients for all the conditions (Diabetes/Hypothyroidism/Hypertension/PCOS)**

**We run each module and scrape recipes for all conditions (Diabetes/Hypothyroidism/Hypertension/PCOS)**

**From given ingredients for allergies, we have split into Allergies commonly in one excel and Nuts Allergies in one excel. (combined with all conditions)**

**Every condition we have added code to check Allergies and Nut Allergies (Diabetes/Hypothyroidism/Hypertension/PCOS)**

**Scraped Recipes Module wise:**

PCOS-54

Diabetes-74

Hypertension-198

Hypothyroidism-49

Allergies-200

Nut Allergies-71

**To extract text from image we have tried different feature using Tesseract**

**We are capturing image from website and trying to save in an image folder and trying to get the text of it.**

A computer code on a white background

Description automatically generated

**Recipe/Food Category:**

**We used common xpath for both so we get all the required details for most of the recipes**

//a[@itemprop='recipeCategory']

A screenshot of a computer

Description automatically generated