**Programming test**

**Instructions**

* Wherever possible enhance the readability of your code by adding comments
* The codes will be tested on a local machine (in addition to your machine) therefore do not hard-code your file system for any input or output.
* If you use any additional data, keep the data in the same folder as your code. Do not make subfolders.
* Title your codes as Prog\_1\_Yourname.py
* You can zip up all your codes in one folder or make separate folders for each question or upload to Github. Please include any README file that you think would be required to run your codes.
* Please mention whether you have used Python2.7 or 3+

**Codes**

1. **API calls**

Raman housings pvt ltd needs to collect some data about houses in a certain area within the US. Can you help them to collect data about atleast 500 homes from the list (House\_features.tsv) that the company has provided us?

Zillow is a real estate online portal which contains information about almost all houses in the US. Fortunately, they provide an API service which we can call to fetch data we need.

Write a python code that uses Zillow’s APIs to fetch the property details about each home (<https://www.zillow.com/howto/api/GetUpdatedPropertyDetails.htm>). We need only the Edited Home Facts, Neighborhood, School district data about the home (see the ‘The results set of the API:’ section). Since you need the zpid to make this call, use their DeepSearch-Results API (<https://www.zillow.com/howto/api/GetDeepSearchResults.htm>) to get that.

**Finally write all the data you collected + the property address into a single csv file (tab delimited).**

**Note:** You would be required to set up your api key (https://www.zillow.com/webservice/Registration.htm). Use <http://www.contata.com/> as the example URL to use the Proper Details API.

1. **RESTful API services**

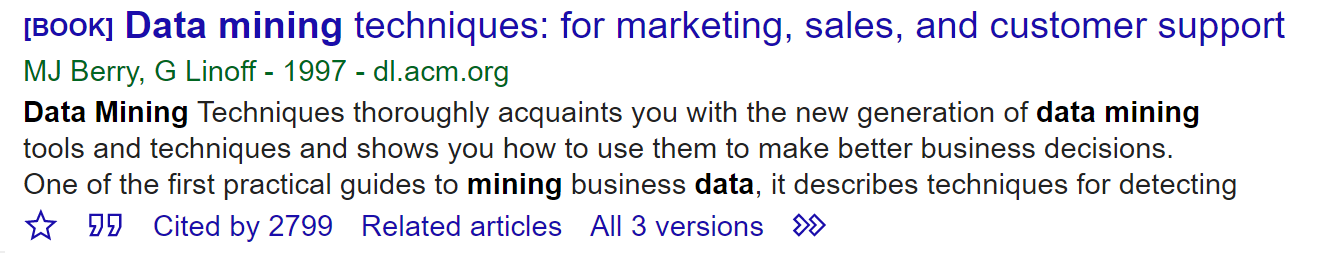
Contata Solutions needs to host a REST API service for its client Raman Photo editing services Pvt Ltd. You have to expose two end points for the client: 1) a POST end point which accepts an image URL and returns whether the image contains borders 2) a POST end point which accepts an image URL and returns the edited image filename i.e. image with borders removed. Since we are first building it in development environment, assume you will send a file name as input to both the end points. Therefore, store the image files in the same folder as the code. And the output of the second end point to the end user is a status saying image edited + edited image file name. Store the edited image in the same folder but with a different name than the original file. Develop the REST API to run on local host. Hint: use flask, check out this image editing service <https://github.com/embali/enimda/> or you can use some of our code (Image\_border\_detect\_crop.py). Sample images are also in the folder.

1. **Web scraping/ Data collection**

Choice 1:

Write a python scraper to collect data from Google scholar for top-10 search results of each of the following queries: 1) data mining 2) machine learning 3) artificial intelligence 4) data analytics 5) text mining 6) image processing. The data collected should contain the article title, author names, year of publication, cited by. Store the data in a tab delimited file with column names as described above including an extra column for the input query.

Title



Cited by

Year of Publication

Authors

Choice 2:

Write a python scraper to collect image data from Google or Bing Images for the following queries: 1) dining room with fireplace 2) bedroom with high ceilings 3) bathroom with natural light 4) kitchen with granite countertops 5) living room 6) entrance foyer. For each query you can collect 20-50 images. For each query store their output images fetched from web into separate folders. The folders can be names with the query itself. Your script should automatically create the folder names when the query is input to your script.