

Winter Workshop on Data Science and Machine Learning
26th - 30th December, 2017

Task Sheet - Day 1

Python:

References:

Lecture Slides:

Tutorial: <https://www.tutorialspoint.com/python/>

1. Take an arithmetic expression as input and solve it.
eg. python arithSolver.py 6+5*4/5
2. Take a bunch of strings as input and output them in alphabetical order.
eg. python strSorter.py igit igdtuw india
3. Take three numbers as input and find largest among them.
eg. python findLarge.py 45 87 12
4. Take two strings as input and find whether second string is present in first string.
eg. python findPattern.py igdtuw gd
5. Use iterative method to find factorial of a given input number.
6. Use recursive method to find factorial of a given number.
7. Write program to illustrate user defined module and its usage.
8. Read an input file containing names, roll numbers and marks of 10 students. Each student's details are present in single line separated by comma. After reading the contents from file, store in a list of dictionary objects variable.
9. Read the variable created above, randomly change marks of students and store them in a file such that each student's details is in one single line.
10. Write a program that creates a divide by zero exception.

Facebook:

Facebook Graph API: <https://developers.facebook.com/docs/graph-api>

API Graph Explorer: <https://developers.facebook.com/tools/explorer/>

Sample Tutorials:

T1.

<https://towardsdatascience.com/how-to-use-facebook-graph-api-and-extract-data-using-python-1839e19d6999>

T2. <http://nodotcom.org/python-facebook-tutorial.html>

1. Create an app as a developer on Facebook to obtain appID. Write a Python program using this appID. Since you would be the 'user' of this app created by you, so you can grant all permissions to this app so that it collects most of your information. Specifically, your python program should output your own feed information and list down 'all' your friends. Do 'all' your friends get listed ? Give reasons for your answer.

2*. Continue with the above Python program, grant specific permissions of your own account to this program (app) and see the output of those permissions in terms of the information that program (app) is able to collect. Present your observations in form of a table whose first

column is name of permission and second column is the information which app (program) was able to obtain after getting the requisite permission.

3*. Create a simple login page using HTML as front end, PHP as server side script and SQL which has user's password stored which is compared with user logs in. Next, modify this login page to provide an option to end user to login using Facebook Login (refer here). Expose the user to a permission dialog and save user's access token (which is long number that encodes what permissions user has granted to your website). Use the access token obtained in the python program written in Q1 to obtain more information about the user from his/her Facebook account.

**: These tasks are optional, do them after completing tasks from all days.*

Twitter:

Twitter API: <https://developer.twitter.com/en/docs>

Python Modules: <https://twython.readthedocs.io/en/latest/> or <http://www.tweepy.org/>

Tutorial: <http://socialmedia-class.org/twittertutorial.html>

1. Write a python program that takes any username or user handle (eg. 'SrBachchan' is username of actor Amitabh Bachchan) as input and collects all tweets posted by that user.

2*. Write a python program that any username or user handle as input and finds usernames of all of its followers. For each follower, find all tweets posted by that follower.

3*. Extend your program in Q2 to store the information in two collections (say A and B) in a mongoDB database.

JSON object structure in collection A:

```
{
input_username: 'xyz'
follower_username: 'abc1'
}
{
input_username: 'xyz'
follower_username: 'abc2'
}
and so on ...
```

JSON object structure in collection B:

```
{
follower_username: 'xyz'
tweet_id: '1234'
}
{
follower_username: 'xyz'
tweet_id: '5678'
}
```

and so on ...

**: These tasks are optional, do them after completing tasks from all days.*

Web Scraping:

References:

Tutorial: BeautifulSoup

<https://www.analyticsvidhya.com/blog/2015/10/beginner-guide-web-scraping-beautiful-soup-python/>

Scrapy: Python Module <https://scrapy.org/>

1. Take a closer look at <https://news.ycombinator.com/> page and observe how information is organized. Think of this information in form of JSON object. Write python program using either beautifulsoup or scrapy module to scrap relevant information and store as JSON object in mongoDB database.

Selenium:

Installation Steps:

- a) Install selenium:- `sudo pip install selenium`
- b) Download Chrome driver or firefox driver using below link :
<https://chromedriver.storage.googleapis.com/index.html?path=2.30/>.
- c) You might need to update Your chrome for it : `sudo apt-get install chromium-browser` or update manually.
- d) Unzip downloaded chrome driver /firefox driver file.Type following commands on terminal:

```
mv chromedriver /usr/bin
sudo chmod a+x chromedrive
```
- e) Install Appium:-

```
sudo apt-get update
sudo apt install nodejs-legacy
sudo apt-get install npm
sudo pip install Appium-Python-Client
```

References:

1. Install chrome driver using following the instructions as per above tutorial. Write a basic program that logs into Facebook using username and password of your own account.
- 2*. Extent the above program to automatically go to the page (using selenium) which shows your friends. Use web scraping method to scrap username of all your friends.
- 3*. Further, automatically click (using selenium) on each of your friend to go to their feed page. Use web scraping to scrap all status updates which your friend has made visible to public and friends.

**: These tasks are optional, do them after completing tasks from all days.*

