CSE648- Privacy and Security on Online Social Media

Midsem Hack

Maximum Marks: 60 **Deadline: 27 Feb 2020**

Instructions :

1. Languages allowed for analysis: Python/Java.
    a. If you are using Python, it will be good to use Jupyter notebook, to show analysis, graphs, and code.
    b. Document your code properly.
2. Make sure you submit .py versions of all the jupyter notebooks.
3. You are free to any online sources, please cite them
4. Please write your own code. All codes will be tested for Plagiarism and if found, institute policy for plagiarism will be followed.
5. You can use any database for storing the data but it will be checked at the time of demo.
6. Write all the analysis along with graphs, charts, etc in **analysis.pdf**
7. Make a readme.txt file with instructions on how to run the code. All libraries, sources, etc used should be properly mentioned in it.
8. Do the hack in the group of at most 3 students.
9. Mention the name and Roll numbers of your group members in all the files attached properly.
10. Zip all your code files along with analysis and readme file in `Rollno_Name_midsem.zip` format. Example `201402230_Swati_midsem.zip`. One submission per group.

---

1. Identify 3 faculties of IIIT Delhi of your choice. Fill your choices here (at least 2 distinct faculties must be chosen, i.e. names that have not yet been filled in the Excel sheet).

2. Gather Personally Identifiable Information (PII) from at least 3 sources on the internet (Consider platforms like Twitter, Facebook, Research Gate, Google Scholar, LinkedIn or any other website faculty is mentioned on.)
    a. Identify the Linked Information of the Faculties.
    b. Identify the Linkable Information of the Faculties.

Definitions:

Linked Information: Any piece of personal information that can be used to identify an individual and includes, but is not limited to Full name, Home address, Email address, Social security number, Passport number, Driver's license number, Credit card numbers, Date of birth, Telephone number, Login details, etc.

Linkable information: Any information that on its own may not be able to identify a person, but when combined with another piece of information could identify, trace, or locate a person. For Example: First or last name (if common), Country, state, city, postcode, Gender, Race, Non-specific age (e.g. 30-40 instead of 30), Job position and workplace, etc.

Note: Gather at least 3 PII of the Faculty which is not present on Faculty's IIITD Website. If not available report along with sources considered.

(30 marks)

3. Gather Faculty's social media posts (as many as allowed by API) and build a timeline of life events like Travels, Conference visits, Awards, Bio change etc.

(10 marks)

4. Identify the people who have the highest engagement with the faculty on social media websites. Report results for different kinds of engagements e.g likes, shares, retweets, etc. Do these people change with time? Analysis for a full history and in blocks of 6 months.

(10 marks)

5. Make a webpage (using a framework of your choice) for each faculty to show all the above-gathered information. (Website can be minimalistic and readable).

(10 marks)

Deliverables: Webpage, Database of collected data, Analysis Report, Source code