INTERACTIVE WEB CRAWLER

Tired from a hectic semester! Want something new? Yes, you're at the right place. A web crawler is like a digital detective that scours the vast expense of the internet, hunting for information hidden within websites. These clever programs, often referred to as "bots", have an important role: they help build the vast catalog of the internet, much like the index in a library. So, are you ready to delve into this fascinating world of information gathering?

Objective:

Our adventure begins with the creation of a web crawler using Python and its powerful libraries. This will recursively extract all the links of a webpage up to a certain user-specified depth. We'll take the data we've gathered and bring it to life with informative and visually engaging plots using Matplotlib. As a grand finale to this interactive journey, we'll empower you to create your very own webpage using HTML. Yes, you read that right! You will not only access the webpage but also run our web crawler directly from it, putting the power of web crawling at your fingertips.

Prerequisites:

There are no hard prerequisites. A basic knowledge of python and its libraries will give you a good head start, but you can even learn them while you start upon your journey.

Skills that you will learn through this project:

- Python (One of the most famous and versatile programming language)
- Numpy (Python library to speed up computations)
- Matplotlib (Used to construct plots and graphs)
- HTML, CSS, Javascript (This combo will be used to make a webpage)
- Other python libraries (To extract websites data)
- Working in a team (Most Important Skill !!!)

Tentative Schedule of the project:

- Week 1: Basics of Python
- Week 2: Introduction to Numpy, Matplotlib and other Python Libraries
- Week 3: Develop Web Crawler
- Week 4: Learn HTML, CSS and Javascript and develop webpage
- Week 5: Integrate Python on backend of the website

Mentors:

- **Saksham Rathi**, sophomore, CSE Department, Contact Number - 6375663539, Email: 22b1003@iitb.ac.in
- **Nandan Manjunath**, sophomore, CSE Department, Contact Number - 7093361624, Email: 22b0920@iitb.ac.in