Soumya Darbha

(470)-652-9522 | sdarbha1@student.gsu.edu
Github: - https://github.com/soumyadarbha | https://soumyadarbha.github.io/
LinkedIn: - https://www.linkedin.com/in/soumyadarbha/
Tableau: - https://public.tableau.com/app/profile/soumya.darbha

Education

Master of Science: Computer Science, Georgia State University	2021-Present	**GPA-3.92**
Bachelor of Technology: Computer Science, JNTU	2013-2017	**GPA-3.9**

Coursework: Adv DB Systems, Data Visualization, Big Data Programming, Cryptography, Python Programming.

TECHNICAL SKILLS

Programming Languages	: Python, Java, C
Data Visualization libraries and Tools	: Numpy, Pandas, Matplotlib, Seaborn, Tableau, MS Excel
Database Technologies	: MySQL, Postgre SQL
Operating Systems & Others	: Windows, Linux, Spark, Github, HTML , MS Word, MS Powerpoint

Professional Experience

Evidence-Based Cybersecurity Research Group, Georgia State University   **Jan 2021 – Present**
Graduate Research Assistant

- Involved in collecting data from 60 underground online darknet sites (markets) and Telegram data (Groups and Channels) that facilitate illicit sales.
- Run queries using **MySQL** to retrieve relevant and most important data. Analyze it using **Tableau** to get information/insights on who is involved, how they operate, and test strategies for disrupting their activities.
- Write **Scrapers** and **Parsers** using **Python** to retrieve the data from darknet sites (markets and forums) and store it in the database.
- Scraping and Parsing data from darknet sites manually takes weeks which is reduced to 1 hour with the use of Scrapers and Parsers code and improved Data efficiency by 20% by generating Dashboards using Tableau and SQL.

Accenture, India   **Jul 2017 – Oct 2020**
Data Analyst

- Worked on a Content Management Project for the Client Hess involved in the exploration and production of crude oil and natural gas. We had 3 applications that stored all the client's oil and gas-related data like cheques, planning documents, SAP, and other financial documents.
- My role involved analyzing user requirements and developing tools accordingly which automated and reduced manual intervention.
- Developed the '**Active User Report tool**' using **Python** and **MySQL** which provided active user's information from the repository and helped to monitor the company's product license. Analyzed and processed this data using **Tableau** by creating dashboards to visualize the data and accurately present the insights of customer data.
- Developed the '**Bulk Documents Export tool**' using **Python** and **MySQL** which helped end-users to extract bulk documents from the repository by reducing the run time from 1 hour to 20 mins.
- Directly worked with clients and customers to resolve the tickets/issues logged by them in **ServiceNow**.
- Retrieved the above customer's tickets data to perform analysis using **Tableau** by creating **Dashboards** to monitor and report the customer's issues in order to automate the issues to reduce manual work and improved efficiency by 40%.

Infosys, India   **Jan 2017 – May 2017**
Intern

- Worked on major project **"Airbnb Big-Data Analytics"** to analyze the top 10 economically best and safe Airbnb stays and report price prediction using top-k and price prediction algorithms. Analyzed the result by creating charts and dashboards in Tableau.
Tools Used: **HDFS, Spark, MySQL, Tableau**.

ACADEMIC PROJECTS

- **Academic Publication Citation Recommendation System:**
  Built Relevance recommendation based on mutual citation i.e list of papers sorted by most count of common cited papers using MySQL, Spark and the concept called Text Similarity.
- **Analysis on FBI Gun Dataset**:
  Analyzed the FBI Gun Dataset using numpy, pandas, and seaborn.
- **Data Analysis on Graduate Admissions**:
  Analyzed the Graduate Admissions Dataset using numpy, pandas, and matplotlib.