Samar Dikshit
Boston, MA | 929-217-0015 | dikshit.s@northeastern.edu
samar14641.github.io | github.com/samar14641 | linkedin.com/in/samar-dikshit

EDUCATION

Northeastern University, Boston, MA — September 2019 – Present
Khoury College of Computer Sciences — Expected Graduation: December 2021
Candidate for Master of Science in Data Science, GPA: 3.78
Related courses: Data Management and Processing, Machine Learning, Information Retrieval, Algorithms

Manipal Institute of Technology, Manipal, India — July 2015 – July 2019
Department of Information and Communication Technology
Bachelor of Technology in Computer and Communication Engineering (Data Analytics minor)

TECHNICAL KNOWLEDGE

Programming Languages:	Python 3, R, C++
Data Science Technologies:	Jupyter, Matplotlib, NetworkX, NumPy, pandas, GeoPandas, PyTorch, scikit-learn, SciPy, Seaborn, Rasa, caret, tidyverse, MySQL, Elasticsearch, Git
Operating Systems:	Windows, Ubuntu

EXPERIENCE

Analytics Engineering Intern — January 2021 – Present
DTonomy Inc., Cambridge, MA

- Analyzing cyberattack data in Python based on the MITRE ATT&CK database to create an algorithm and API to generate a graph of attack patterns for the SOAR platform
- Creating bots on Slack using Rasa 2 to use services like Google Analytics and AbuseIPDB
- Developing Elastic Security integrations for the SOAR platform using Node-RED

Research Assistant — June 2020 – January 2021
Center for Complex Network Research, Northeastern University, Boston, MA

- Collected, processed, and explored data with over 1.5 million samples related to philanthropies, non-profits, and universities from the LittleSis API using Python
- Worked on matching the names of organizations and people across the new data and previous GuideStar data using TfidfVectorizer, CountVectorizer, and pairwise kernels, thereby expanding the previous network
- Created and analysed a new network with more than 97,000 relationships between sociopolitical entities to determine factors that influence grants and donations using NetworkX and pandas

Teaching Assistant — May 2020 – December 2020
Northeastern University, Boston, MA

- DS2000 Programming with Data, CS3000 Algorithms and Data

PROJECTS

COVID-19 Visualizations (non-academic project) — April 2021

- Created visualizations in Python using Matplotlib to track the progression of, and statistics related to the COVID-19 pandemic in Massachusetts
- All graphs were created using colourblind-friendly palettes to ensure accessibility for individuals with protanopia, deuteranopia, and tritanopia

Detecting Brain Tumours using Machine Learning — October 2020 – December 2020

- Trained a set of classifiers to detect a brain tumour when given an MRI scan in Python
- Used decision trees, adaptive boosting, and a convolutional neural network to obtain a peak sensitivity and accuracy of 98.27% and 99.17% with cross-validation, hyperparameter tuning, and feature selection

Similarities and Differences between News Sources in the United States — October 2019 – November 2019

- Developed a set of filters in R to obtain articles related to politics out of over 72,000 articles scraped from various news websites such as BBC, CNN, Fox, and FiveThirtyEight
- Used visualizations in R illustrating word associations, sentiment, and bigrams to prove the existence of media bias across different sources based on political leaning

More projects can be found here.