

# SHAMIKH HOSSAIN

(469) 386-1378 ◊ shamikh.hossain@duke.edu ◊ shamystic.github.io ◊ github.com/shamystic ◊ linkedin.com/in/shamikh

## EXPERIENCE

### Valassis Digital (formerly MaxPoint)

Software Engineer Intern, Platform & Services Team

Jun 2018 - Present

Austin, TX

- Working to optimize distributed dataset creation by porting ETL job from Java/Cascading stack to Scala/Apache Spark for real-time bidding system's text-classification pipeline. Ensuring backward compatibility, test coverage using ScalaTest suite, and analyzing performance improvements in production cluster using Graphite/Grafana monitoring platforms.

### Duke Energy Data Analytics Lab

Undergraduate Researcher, Data+ Summer Fellow

May 2017 - May 2018

Durham, NC

- Developed techniques for predicting village-level household electrification rates from satellite imagery using supervised learning
- Curated ground-truth dataset, designed and implemented end-to-end machine learning pipeline to streamline feature extraction, selection, and dimensionality reduction using Python, pandas, scikit-learn, matplotlib and Jupyter notebooks
- Applied gradient boosting classification to predict electrification with .80 AUC using class-imbalanced, high-dimensional dataset

### Duke Office of Information Technology: Innovation Co-Lab

Developer, Technical Consultant

Jan 2018 - Present

Durham, NC

- Develop and improve Ruby on Rails homepage web application and project discovery platform for students: colab.duke.edu
- Build and maintain Co-Lab infrastructure and developer tools, hold technical advising office hours on full-stack development

### Carin Lab Group at Duke University

Undergraduate Researcher

Nov 2017 - Present

Durham, NC

- Applying semi-supervised deep learning techniques to Duke Hospital's medical datasets, advised by Dr. Lawrence Carin
- Implemented an autoencoding NLP framework using TensorFlow for parsing diagnosis information from medical notes

### Duke Office of Information Technology: Biology

IT Consultant

Aug 2016 - May 2017

Durham, NC

- Worked alongside full-time IT analysts to provide tailored hardware and software support for professors, staff, students
- Wrote documentation, trained new consultants, implemented department-wide software integrations and security updates

## EDUCATION

### Duke University

B.S. in Computer Science, B.S. in Statistics, Minor in Economics

2016 - 2020

**Coursework:** Probability, Microeconomics, Data Structures, Linear Algebra, International Trade, Discrete Mathematics

## SELECT PROJECTS

**SFPD Dispatch Analysis:** Investigation of San Francisco emergency dispatch data to formulate solutions for improving response times. Recognized as top entry to Capital One MindSumo CS challenge. (Python, matplotlib, HTML/CSS/JS)

**Kickback:** Personal finance web application built with Visa and Capital One APIs that uses transaction data to identify spending habits and recommend relevant discounts. Winner of Capital One SWE Summit Hackathon. (Flask, HTML/CSS/JS)

**HitChecker:** Review tool for crowdsourced image annotation - drastically reduced time needed for dataset review/worker payment and ensured integrity of labeled data for Data+ research project. (Tkinter, Amazon Mechanical Turk API)

**Politiquette:** Chrome extension displaying on-hover interest group ratings for US senators on inequality issues (Flask, JS)

## SKILLS

**Software:** Python (proficient); Java, Scala, C, Ruby (prior experience); Git

**Data:** NumPy, SciPy, pandas, matplotlib, Matlab, scikit-learn, R, Jupyter Notebook, Tableau, Excel

**Web:** HTML/CSS, JavaScript, Ruby on Rails, Django, Flask

**Coursework:** *Enthought Training:* SciPy & NumPy, *DeepLearning.ai:* Neural Networks & Deep Learning

## OTHER

**Activities:** Planned and organized Duke Machine Learning Day, Duke's first workshop/research symposium event for ML Director of Finance of Duke Undergraduate Machine Learning, organizing on-campus speaker events and fall datathon

**Interests:** Statistical modeling, API development, machine learning, distributed computing, high-dimensional data analysis

**Recognition:** IB Diploma Recipient, National AP Scholar (2x), U.S. Presidential Scholar Candidate, Burger King Scholar, Dean's List, Microsoft Coding Competition - 3rd Place (tie), Duke Research Computing Symposium Poster Competition - 1st Place, 2018 State Energy Conference of North Carolina - Honorable Mention Research Poster

**Events:** Attended Capital One Software Engineering Summit (May 2018), AQR (Applied Quantitative Research) Capital Management Early Engagement Program (May 2018)

**Work Authorization:** U.S. Citizen