Malone Center for Engineering in Healthcare Data Science Internship

Position Summary (to openings): The data science intern will be responsible for modeling complex problems, discovering insights and identifying process improvement opportunities through the use of statistical, mathematical, computational techniques. In addition to analytical skills, the data science intern should be proficient at integrating and preparing large, varied datasets, and communicating results. The data science intern will work closely with the Johns Hopkins Operations Integration team and clients, data stewards, project/program managers, and other IT teams to turn data into critical information and knowledge that can be used to make sound operational/clinical decisions. Other responsibilities will include providing data that is congruent and reliable and, the need to be creative thinkers and propose innovative ways to look at problems by using data mining and mathematical modeling on the set of information available.

The data science internship requires a full-time commitment. It is not permissible to hold other employment during the internship. There are no vacations during the program.

Benefits for the Student: Acquire theoretical and practical training in advancing the organizational performance of The Johns Hopkins Health System using data science and systems engineering. The student can expect an experience similar to a lab rotation, as well as be out in the hospital floor to become acquainted with the assigned project and the operational and clinical workflows. Before arrival, each data science intern receives multiple papers related to their assigned project. The goal of the student's project and its relationship to other work in the area will be discussed, and the student will be trained in the techniques necessary to conduct data and systems analyses.

The data science intern will be part of Operations Integration, which is an in-house consultancy group providing services to different areas of the Johns Hopkins Hospital in pursuit of process improvement opportunities. Operations Integration uses data-enabled science to re-engineer healthcare delivery. This means creating systems that transform raw data into information providing evidence for medical and operational decisions, organizational strategies, and policy-making.

Compensation: Approximately $5,000, depending on schedule and duration.

Required Education: Bachelor's degree in mathematics, statistics, or computer science. Master's students preferred although strong undergraduate students will be considered (undergraduate students must have completed at least 2 years of college).

Preferred: Demonstrated interest and potential to pursue a Ph.D. degree.

Required Experience: No previous research/industry experience required.

Desired knowledge, skills, and abilities:

- Solid knowledge of statistical and mathematical modeling techniques
- The ability to come up with solutions to loosely defined business problems by leveraging pattern detection over large datasets
- Strong programming skills (R, SQL, Arena, GAMS)
- Strong background (optimization, network models in operations research, Markov chains, machine learning, Monte Carlo methods, discrete-event simulation)
- Must demonstrate strong critical thinking and analytical reasoning skills

- Ability to execute assigned project tasks within established schedule
- Possesses sound documentation skills; writes and communicates clearly and concisely

Application Process: Send your CV, cover letter (describing relevant course work, research experience, and/or future plans about industry/research career) and one letter of recommendation/professional reference contact info.

Email applications to Vess Vassileva-Clarke: vclarke@jhu.edu
Subject line: MCEH Internship application

Application deadline: July 1, 2017

Anticipated start date: July 10, 2017, duration appx 10 weeks

Short project descriptions:

Project #1	Using Machine Learning and Dynamic Social Networks for Information Extraction from Hospital Real Time Location Systems
Description	Real time location systems contain massive amounts of timestamps data about the location of life-saving assets throughout the hospital. However, the large amount of data makes this information hard to extract automatically. In this project, we will use machine learning and dynamic social networks to develop a pipeline for automatically extracting information from location systems that could later be used in optimal allocation of medical equipment.
Tasks	- Weeks 1 & 2: Scoping the Problem—literature review (methods and application area) as well as visits to the hospital floor to understand the process - Week 3, 4 & 5: Exploring and Preprocessing Data - Week 6 & 7: Data Mining and Pattern Evaluation - Week 8 - 10: Data Solution Deployment
Recommended Skills	Good background in programming (at least one class in object oriented programming and/or familiarity with R). While useful, a medical background is not required. Strong background in statistics.
Start Date	Summer 2017
Duration of Internship	Maximum 10 weeks

Project #2	Using Time Series Modeling for Evaluation of Process Improvement Efforts
Description	New operational and policy interventions are being carried out to improve hospital operations. Evaluation of the impact of these interventions is not only essential for improving patient care, but also the key to sustained improvements and evidence-based medicine. In this project, we will use interrupted time series design to estimate the effect of time-delimited interventions in non-randomized settings.
Tasks	• Weeks 1: Scoping the Problem—literature review (methods and application area) as well as visits to the hospital floor to understand the process • Week 2 & 3: Collecting additional information and interviewing key process improvement champions • Week 4 & 5: Exploring and Preprocessing Data • Week 6 & 7: Statistical Analysis and Pattern Discovery • Week 8 - 10: Documenting and reporting results to leadership
Recommended Skills	Good background in programming (at least one class in object oriented programming and/or familiarity with R). While useful, a medical background is not required. Strong background in statistics. Strong presentation skills.
Start Date	Summer 2017
Duration of Internship	Maximum 10 weeks