Data Scientist

The Li Ka Shing Centre for Healthcare Analytics Research and Training (LKS-CHART) is looking for a Data Scientist who will focus on machine learning applications in healthcare.

The Li Ka Shing Centre for Healthcare Analytics Research and Training (LKS-CHART) of St. Michael's Hospital is an exciting new centre that houses a team of Data Scientists devoted to applying data science to the healthcare field to advance patient care and health system efficiency. The LKS-CHART works closely with physicians and administrators at St. Michael's Hospital and researchers at the University of Toronto in a team-based environment. The Data Scientist will use the Enterprise Data Warehouse at St. Michael's Hospital as the primary data source to provide data-driven insights into key drivers of patient outcomes and hospital performance. The Enterprise Data Warehouse at St. Michael's Hospital is one of the most comprehensive inpatient databases in Canada.

The main roles of a Data Scientist (called a Senior Data Science Specialist at St. Michael's Hospital) are to design scientifically rigorous approaches to examining data, apply advanced analytical techniques to healthcare data, and translate the findings into meaningful, applied knowledge for end users. This includes conceptualizing algorithms and study designs, creating statistical analysis plans, conducting and interpreting analyses, and reporting findings in an applied manner. The Data Scientist role is a service-based role, rather than a traditional academic research role. However, depending on the interests of the individual, some level of academic engagement can be explored. Each Data Scientist at the LKS-CHART has an area of specialization, such as machine learning, simulation modeling, biostatistics and financial analytics. *The area of specialization for this Data Scientist position is computational statistics with a focus on machine learning. This is a mid to senior level position and the rank of the individual will be determined based on their qualifications and experience at the time of the hiring.*

MAIN DUTIES & RESPONSIBILITIES:

- Develop research study plans outlining key components of analytical approaches. Provide on-demand consultative data science expertise for ad-hoc requests and recommend data science approaches to meet client needs
- Work closely with potential clients to establish coding techniques, structure of dataset for studies and develop sound analytical plans
- Carry out machine learning tasks that involve supervised learning, unsupervised learning, and reinforcement learning
- Conduct analyses using machine learning approaches such as decision tree learning, association rule learning, artificial neural networks, Bayesian networks, support vector machines, clustering, representation learning, and deep learning.
- Ensure adequate quality control by setting standards, monitor results and institute appropriate steps for data cleaning, consistency checks and other data quality control measures prior to analysis
- Carry out all computer programming required to execute analyses and develop efficient standardized programs, algorithms, or systems to enhance efficiency
- Maintain data documentation, physical and logical storage of scripts, records and master archive lists
- Critically appraise/edit code of trainees, students etc.

- Integrate with the other Data Scientists at the LKS-CHART to develop multi-disciplinary analytic approaches, and assist in best practices and process development
- Analyze/interpret results of research studies – draw conclusions for review and incorporation into reports/presentations
- Prepare output of analyses for review including graphs, tables, and interpretation
- Ensure accuracy of all analyses and appropriate interpretation in study reports
- Collaborate with clients, students and other Data Scientists on academic presentations and co-authorship of publications
- Critically appraise/edit manuscripts of trainees, students etc.

QUALITIFICATIONS

- An undergraduate degree in Computer Science with significant course work in machine learning is required.
- A Master's degree in Computer Science with a focus on machine learning is strongly preferred.
- A PhD degree in Computer Science with a focus on machine learning is highly desirable.
- Experience working with large health datasets is preferred
- For senior candidates, at least 5 years relevant experience in research or in industry
- Mastery of R and Python; mastery of other statistical programming languages an asset
- Demonstrated mastery of computational statistics methods with a focus on machine learning with an ability to develop and code algorithms over and beyond the standard ML tools and libraries
- Deep mathematical understanding of and experience with machine learning approaches such as decision tree learning, association rule learning, artificial neural networks, Bayesian networks, support vector machines, clustering, representation learning, and deep learning
- Experience visualizing data for a broad range of audiences an asset
- Data management and/or monitoring experience an asset