Institute for Big Data Analytics

To create knowledge and expertise in the field of Big Data Analytics by facilitating fundamental, interdisciplinary and collaborative research, advanced applications, advanced training and partnerships with industry.

Research - Training - Outreach

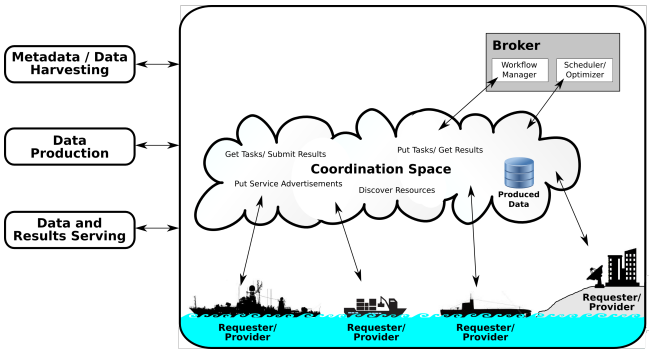
Newsletter

Spring 2017

Research Project in Support of the Royal Canadian Navy

Since the autum the Institute for Big Data Analytics has been involved in a challenging project with the Department of National Defence and Lockheed Martin Canada, entitled “Mission-Relevant Information Management for Integrated Response" (MIMIR). This three-year project is attempting to study the real-time information needs of decision makers aboard Royal Canadian Navy vessels and to provide tools for managing the increasing volume of data that feeds into those decision processes. There are many challenges to supporting these needs in a dynamic and unpredicatable environment, requiring a multi-disciplinary team. From Dalhousie Stan Matwin represents the Institute for Big Data Analytics and brings his expertise in Data Science, including data analytics, data visualization and data management. Andrew Rau Chaplin contributes his experience in high performance computing, analytics, risk management and cloud computing. And Derek Reilly contributes his experience in human-computer interaction and visualization. From Defence Research and Development Canada Anthony Isenor is the leader of the Maritime Information Support Group and Sean Webb is a computer scientist with that same group. From Lockheed Martin the collaborators are Maybelline Lee and Gabi Salameh. All of this expertise will be called on in the goal of designing tools and processes to support the information needs of the future.

This applied research is expected to produce a suite of processes and prototypes and to provide research training opportunities along the way to a number of graduate students and postdoctoral fellows. Funding provided by NSERC, DRDC and Lockheed Martin.



TRIBE Student does Internship at PlaceSpeak Inc

An important part of the graduate training program in Big Text Data (TRIBE) is an industrial internship where the student gets the opportunity to work in a professional environment and to apply their skills to focused, industrially-relevant research projects. Masters Student Lulu Huang recently completed her internship with Vancouver based company PlaceSpeak Inc. PlaceSpeak is a location-based civic engagement platform whose mission is to build legitimacy in online democratic practices by authenticating digital identity to place, protecting individual privacy, and closing the feedback loop between public consultation and accountability. The system allows online feedback with a guaranteed local audience.

The challenge for Lulu was that this online participation produces large quantities of text data. She was not given an exact task but had been invited to follow her own instincts and to use her expertise to find value in the data. This involved initially a lot of reading to try to discern the hot topics in a given issue. Sentiment analysis had some applications to the data and visualization was tricky as there were many different ways to do map-based visualizations.

Lulu is enjoyed the internship, especially the opportunity to get some experience about how a company is run. The discipline of daily reporting was a little different than the academic environment. Lulu also benefited from her stay in Vancouver by connecting and consulting with Dr. Fred Popowich from SFU, one of the TRIBE CREATE partners and co-applicants. Lulu has gone on to finish her Masters Degree and is now living in Vancouver.

DALHOUSIE UNIVERSITY
INSTITUTE FOR BIG DATA ANALYTICS

4th Floor - Faculty of Computer Science - Dalhousie University - 6050 University Ave - PO Box 15000 - Halifax - NS - Canada - B3H 4R2
902-494-8087 bigdata@cs.dal.ca bigdata.cs.dal.ca

News

Student Profile: Hossein Sarshar

Since February, recent graduate Hossein Sarshar (also known as "Parsa") has been working at his dream job of "Data Scientist" for Mobivity, a company based in Chandler, Arizona with an office in Halifax, which focuses on digital marketing for major brands. Hossein is enthusiastic about applying his skills in Artificial Intelligence (AI) to enable marketing approaches to be more effective and intelligent. This is a great step on a career that he has been working towards for the past few years.

Hossein had done his undergraduate degree in computer science in Malaysia and was working in the software engineering industry for a decade but decided to make a transition to the Data Science world after getting inspired by an article he read in 2013 that was titled "Data Scientist: The Sexiest Job of the 21st Century". When he started to think of a change in direction, he considered going to Germany or coming here to Dalhousie. He was eventually swayed towards Nova Scotia, in part because his friend had applied to Dalhousie and in part because he had contacted Dr. Andrew Rau-Chaplin and Dr. Stan Matwin and was enthusiastic about the research that they were doing and wanted to be a part of it. In January 2015, he moved to Halifax with his wife Mona from Iran where he born and raised to start his Masters in Computer Science with a focus on machine learning and big data analytics. He was awarded a scholarship to support his studies under the TRIBE program and became a valued part of the Institute for Big Data Analytics.

Part of his studies involved projects with industry partners. One small project took him to St. John's to work with startup Sentinel Alert. His major focus, however was on a larger project and an internship with SolutionInc, a company providing internet access via both Wi-Fi and wired solutions in high demand areas such as airports, hotels, restaurants etc. The work on this project was an important contribution to his Masters thesis and two published and work-in-progress research papers. Hossein's thesis was titled, "Analyzing Large Scale Wi-Fi Data Using Supervised and Unsupervised Learning Techniques".

And if he wasn't busy enough with these projects he also found time to take part as a team member in two intense competitions, winning first prize in the three-month marathon "Smart Energy App Challenge", and also winning the award of "Most Innovative Use of Big Data" in the coast to coast "Sports Hack" competition. After his success with the Smart Energy competition, Hossein was invited to work as a Data Scientist consultant with the NSCC Applied Energy Research (AER) Lab for the last few months of his studies.

During the two years of his studies at the Institute for Big Data Analytics, Hossein's schedule may have been kind of crazy, but as he told us, "It was one of the most exciting, challenging and rewarding times in my life."

Data Challenges of Property Valuation

Recently, two Masters students with the Institute for Big Data Analytics, Balachandhar Tirunelveli Nallasivan and Thomas Crowell, conducted applied research with Property Valuation Services Corporation (PVSC) to develop new tools to help PVSC manage their large data sets.

PVSC is an independent, not-for-profit organization funded by Nova Scotia's 51 municipalities and is responsible for valuing all real property in Nova Scotia, approximately 600,000 properties, on an annual basis. Each January, PVSC provides an assessment roll to municipalities, which is used to calculate property taxes, and delivers property assessment notices to every Nova Scotia property owner. PVSC collects, reviews and analyzes market sales and property information on an ongoing basis to assist in the determination of property assessments. The organization has a lot of data!

Bala and Thomas worked on two projects. The first was to build an Automatic Valuation Model (AVM) for condo properties, and the second was to explore the current state and impact of short term, non-lease rental properties through applications such as Airbnb.

Machine Learning algorithms, techniques of Convex Optimization and programming languages of Java, Python and R were some of the tools employed by Bala and Thomas on these projects. They are excited that PVSC is pursuing technology solutions within the assessment business. These research projects are expected to continue.

Getting Involved with Local Industry

Recently, the Institute for Big Data Analytics pitched a proposal to the local business community. Called "Got Data? Get Growing" It was an invitation to send us proposals for interesting data problems that our students could get their teeth into and work on as a summer internship. This initiative was made possible thanks to funding from the Big Data For Productivity Congress. The organizers of the Congress had put together a huge meeting in Halifax last year and thanks to prudent management had ended up with surplus funds and a mandate to spend this surplus on good data science projects in Nova Scotia.

We received quite a few interesting proposals from a wide range of organizations both in the public and the private sector. Last summer months our Masters and PhD students, under the supervision of Director Dr. Stan Matwin, worked on helping ViewPoint, the portal for information about properties for sale, to analyse the massive amounts of data concerning those who visit the site: both those who only browse and those who actually make a purchase. They also helped Symphony Nova Scotia to make sense of years of data concerning concert attendees, subscribes and donors and looking for the valuable correlations. Other organizations also with successful proposals were Bell Alliant, Pantel Inc, SolutionsInc Ltd and the Nova Scotia Health Authority with data on alcohol related violence. Internships are an important part of the TRIBE program funded by NSERC supporting the training of students in data science at Dalhousie. We also value the opportunity to make a difference to organizations in our local community.

Researcher Profile: Renata Dividino

In March of this year the Institute for Big Data Analytics was pleased to welcome Dr. Renata Dividino onto the team. Originally from São Paulo, Renata has travelled quite extensively for her studies, finishing her undergraduate degree in Computer Sciences in France and Brazil in joint cooperation between Unicamp, Brazil and the Ecole Central de Lyon, France and going on to do her Masters in Computer Science in the University of Saarlandes and her PhD in the University of Koblenz, both in Germany. Renata has been working in many different research groups in the field of Web Sciences and Semantic Web in Europe, and at the Institute Renata is aiming to learn about more about machine learning, hopefully finding productive ways to combine these fields.

Her first focus with the Institute will be to work on an ongoing project with the Department of National Defence which attempts to learn how to better support the current and future information needs of naval decision makers aboard Royal Canadian Navy vessels. She is also looking forward to bringing her skills to the upcoming MERIDIAN initiative (Marine Environment Research Infrastructure for Data Integration and Application Network).

As well as the challenge of the research Renata is enjoying life in Halifax. She says she is impressed with the kindness and thoughtfulness that Canadian have showed and living here is a good match for her family since people here have the openness of Brazilians and the organizational skills of the Germans. She hopes that her children will learn all that they can from the Canadian experience.

Newsletter Editor:
David Langstroth
dll@cs.dal.ca

"There were 5 exabytes of information created between the dawn of civilization through 2003, but that much information is now created every 2 days."
- Eric Schmidt, Google