Master of Science in Data Science (MScDS)

Starting Date: September 2020

The amount of data collected from sensors, mobile devices, social networks, and communication networks in our world has been exploding, and the speed of growth will continue to expand in exponential fashion. This phenomenon is generally known as Big Data. Data science is the discipline of studying Big Data, involving using mathematic and algorithmic techniques to solve some of the most analytically complex data problems, leveraging troves of raw information to figure out hidden insight that lies beneath the surface of data. It centers around evidence-based analytical rigor and building robust decision capabilities. The goal of the proposed MScDS program is to provide students with solid knowledge, techniques and experience in the research and applications of data science, and to prepare our students to become researchers and leaders in the emerging area of applied research and application at the regional, national and global levels.

The program builds upon the existing strengths of the Departments of Mathematics and Statistics, Computing Science, the Centre of Optimization and Data Science (CODS), as well as the departments' and centre's collaboration with other units across campus and industries, and offers students the opportunity to engage in applied research and applications of data science through interdisciplinary classroom experiences, cutting edge techniques and algorithms, and participation in industrial projects motivated by practical problems. Upon completion of the program, students will have gained key knowledge and skills in applied data science and optimization.

Contact:

Dr. Roger Yu, Coordinator
Email: MScDS-coordinator@tru.ca, Phone: (250) 371-5552

Job market and potential areas of employment for graduates

The demand for people with modelling and data analytics skills is very high at the present and will be even greater in the near future. In a general sense, it is estimated by Canada's Big Data Consortium that by 2018 Canada's overall shortage of professionals with deep analytical skills will be in the range of 14,000 to 19,000, and shortage of data literate managers and analysts will be about 150,000. Employment opportunities for graduates are in the areas of:

- **Data scientists in large private or public organizations**
- **Data analyst positions in software development, business Intelligence, medical/health research, and data management**
- **Strategic analysis manager positions of policy making bodies of government and organizations**
- **Data architect designers of social networks, communication networks and the like**

Program Learning Outcomes

Upon successful completion of the MScDS, students will be able to:

- **Master essential data investigation tools, including data munging, cleaning, sampling, management, exploratory analysis.**
- **Use data exploratory methods to visualize high dimensional data in order to identify the trends and patterns in data sets.**
- **Implement foundational concepts of data computation, such as data structure, algorithms, simulation, and analysis.**
- **Integrate the techniques from mathematical modeling, optimization, machine learning, data mining and applied statistics to model big data set to a workable frame for further data investigation.**
- **Apply the advance data analytical methods and algorithms, including, regression, clustering, classification, prediction, and data communication to large data sets to extract meaningful insights.**
- **Design an analytic strategy to model a potential issue and knowledge-based solution for real-world challenges using public and private data sources.**
- **Ability to communicate results of data analysis effectively (visually and verbally) to a broad audience.**

THOMPSON RIVERS UNIVERSITY

Program Structure:

The MScDS program at TRU consists of *four core courses* plus a choice of one of two completion options: thesis option and graduate project option. With full-time study, the MScDS program is designed to be completed in *four* academic terms.

Each student completes the program with a minimum of 32 credits. Among them,

All four core courses:
STAT 5310, STAT 5320, DASC 5410, DASC 5420
Two terms of graduate seminars:
DASC 6810
Graduation Project or Thesis:
DASC 6910 or DASC 6930
The remaining credits can be any of courses MATH 5210, MATH 5220 or DASC 6XXX not listed above.

Option 1 (Project-based):

DASC 6810 (Graduate Seminar) for two terms	1×2
DASC 6910 (Project)	9
DASC core courses	12
Elective courses	9
Total:	32

Option 2 (Thesis-based):

DASC 6810 (Graduate Seminar) for two terms	1×2
DASC 6930 (Thesis)	12
DASC core courses	12
Elective courses	6
Total:	32

MScDS Core

STAT 5310 Statistical Design and Inference for Data Science
STAT 5320 Linear Models for Data Science
DASC 5410 Data and Database Management for Data Science
DASC 5420 Theoretical Machine Learning

Additional Courses and Credits

MATH 5210 Advanced Modelling Techniques
MATH 5220 Advanced Optimization Methods
DASC 6510 Selected Topics in Data Science
DASC 6520 Directed Studies in Data Science
DASC 6210 Data Analysis in Business and Economics
DASC 6310 Data Analysis in Biology and Life Science
DASC 6710 Work Experience Credits

DASC 6810 Seminar Series

DASC 6910 Graduation Project
DASC 6930 Thesis

Admission Requirements

The admitted students must meet *each* of the following:

1. **Education Requirement** – Acceptable four (4) year bachelor degree in a discipline of science, or a related discipline, with a minimum B average (GPA of 3.00 on a scale of 4.33) for Project Option students and B+ average (GPA of 3.33 on a scale of 4.33) for Thesis Option students, in the last 60 credits.
2. Prospective students are expected to demonstrate working knowledge of statistics, data structures and algorithms, databases and R/Python software packages. These prerequisites are equivalent to the following courses that are offered at Thompson Rivers University: MATH 2110 (Calculus III), MTH-2120 (Linear Algebra), STAT 2000 (Introduction to Statistics), COMP 1231 (Computer Programming II).
3. **Language Requirement** – Applicants must have adequate language training. For students who did not complete their undergraduate degree in an English language university, they need to demonstrate the language competence (e.g., TOEFL score 570 or iBT score 88 or IELTS 6.5 or higher).
4. Two letters of reference from academics or professionals.

Program strengths

- Highly qualified faculty with extensive publications and strong record of prestigious governmental research funding and industrial funding
- Proven record of industrial and community-based joint research projects
- Wide range of expertise among faculty members in data science, machine learning, computational statistics, bioinformatics and discrete mathematics.
- Affiliated with the Centre for Optimization and Data Science, which comes with a membership consisting of scholars and experts from many different disciplines and close connections with local companies and organizations
- Flexibility in program focus (thesis versus project track) based on students' planning of career directions and paths
- Broad applicability of its educational outcomes to sectors throughout software development, data administration, data analysis and business intelligence.

THOMPSON RIVERS UNIVERSITY