**APS1070: Foundations of Data Analytics and Machine Learning (Section 0201)**

| | |
|---|---|
| **Instructor:** | Dr. Jason Riordon, RS214, jason.riordon@utoronto.ca |

**Lecture schedule:**    Section 0201:    Mondays, 16:00-19:00 in GB221, starting Sep 9

**Practical schedule:**
- Section A:    Wednesdays, 15:00-17:00 in GB144
- Section B:    Wednesdays, 15:00-17:00 in GB150
- Section C:    Thursdays, 13:00-15:00 in SF1013
- Section D:    Fridays, 10:00-12:00 in SF1013

**TA contact info:**
- Zheda Mai - zheda.mai@mail.utoronto.ca (Lab Sections A & C)
- Jihwan Jeong - jhjeong@mie.utoronto.ca (Lab Sections A & C)
- Shiva Akbari - shiva.akbari@mail.utoronto.ca (Lab Sections B & D)
- Ali Hadi Zadeh - a.hadizadeh@mail.utoronto.ca (Lab Sections B & D)

**Course description:**

This course covers topics fundamental to data analytics and machine learning, including an introduction to Python and common packages, probability and statistics, matrix representations and fundamental linear algebra operations, basic algorithms and data structures, discrete math and continuous optimization. The course is structured with both weekly lectures (fundamental principles) and labs (practical Python applications).

**Grading:**

| Assignment/Exam | Weight (%) | Due Date / Time |
|:---:|:---:|:---:|
| **Project 1** | 12.5 | Oct 2 @ 12:00 |
| **Project 2** | 12.5 | Oct 16 @ 12:00 |
| **Midterm** | 20 | Oct 26 @ 13:00 |
| **Project 3** | 12.5 | Nov 6 @ 12:00 |
| **Project 4** | 12.5 | Dec 4 @ 12:00 |
| **Exam** | 30 | Dec 10 @ 13:00 |

Project submissions will be online through colaboratory/github. ***It is the student's responsibility to verify that projects are submitted.*** Projects that are late will incur a mark of zero.

**Academic honesty:**

Do not submit code that you have not written yourself. Students suspected of plagiarism on a project, midterm or exam will be referred to the department for formal discipline for breaches of the Student Code of Conduct.

**Student responsibilities:**

It is the student's responsibility to attend lectures and labs, and ensure labs are submitted on time.

**Preliminary schedule of lecture topics and labs:**

| | Wk | Date | Lecture | Book & Chapter* | Lab |
|---|---|---|---|---|---|
| **Python Programming** | 1 | Sep 9 | Course Overview, Python Language Basics, Ipython, and Jupyter, Data Structures, Functions | PFDA 1-3 | Project 1: Basic Data Science |
| | 2 | Sep 16 | SciPy, NumPy, pandas, Data Loading/Storage/File Formats, Nearest neighbour classifier | PFDA 4-6 | Project 1: Basic Data Science |
| | 3 | Sep 23 | Data Cleaning/Preparation, Data Wrangling (Sorting/Searching), Plotting and Visualization | PFDA 7-9 | Project 1: Basic Data Science |
| **Mathematical Foundations** | 4 | Sep 30 | Linear Algebra | MML 1-2, MC 1-3 | Project 2: Anomaly Detection |
| | 5 | Oct 7 | Analytical Geometry | MML 3 | Project 2: Anomaly Detection |
| | 6 | | *No class – Thanksgiving* | | Project 3: Linear Regression |
| | 7 | Oct 21 | Matrix Decompositions | MML 4, MC 4 | Project 3: Linear Regression |
| | | | *Midterm Exam on Saturday, Oct 26, 13:00-14:30 at Exam Centre* | | |
| | 8 | Oct 28 | Vector Calculus | MML 5 | Project 3: Linear Regression |
| | 9 | Nov 4 | Probability and Distributions | MML 6, MC 5 | **NO LAB** |
| | 10 | Nov 11 | Continuous Optimization | MML 7 | Project 4: PCA |
| **Machine Learning** | 11 | Nov 18 | Maximum Likelihood, Regularization, Overfitting, Hyperparameter Tuning | MML 8, MC 6 | Project 4: PCA |
| | 12 | Nov 25 | Linear regression | MML 9 | Project 4: PCA |
| | 13 | Dec 2 | Dimensionality Reduction with PCA | MML 10 | **NO LAB** |
| | | | *Final Exam on Tuesday, Dec 10, 13:00-15:00 at Exam Centre* | | |

*Reference material and chapters, with PFDA = Python for Data Analysis, 2nd Edition, MML = Mathematics for Machine Learning, MC = Matrix Cookbook