University of Montana

ScholarWorks at University of Montana

University of Montana Course Syllabi

Open Educational Resources (OER)

Fall 9-1-2020

M 561.B01: Advanced Data Science Analytics

avier Perez-Alvaro
University of Montana, Missoula, javier.perez-alvaro@mso.umt.edu

Follow this and additional works at: https://scholarworks.umt.edu/syllabi

Let us know how access to this document benefits you.

ecommended Citation

Perez-Alvaro, Javier, "M 561.B01: Advanced Data Science Analytics" (2020). *University of Montana Course Syllabi*. 11370.
https://scholarworks.umt.edu/syllabi/11370

M461 - Data Science Analytics
M561 - Advanced Data Science Analytics
(Fall 2020)

Instructor information:

Instructor: Javier Pérez-Álvaro
Office: M301
Email: javier.perez-alvaro@mso.umt.edu
Office hours: See http://www.umt.edu/people/perezalvaro for up-to-date OH.

Course Format:
Meetings: Monday, Wednesday, Friday 9:00-9:50 a.m.

Course content:
Algorithms are the machinery behind the data analytics (the subject matter and focal point of the course). To be good at data analytics, one must be a competent programmer and have experience with the data and the algorithms of data science. To gain an understanding of algorithm design and good programming practices, students will work with a set of prototypical algorithms that are representative of data analytics. To learn how to function as a data scientist in a relatively short time, the student will be actively engaged in turning the algorithms into code and using them with real data

After successfully completing the course, you will have a good understanding of the following topics and their applications:

1. Data Visualization, Data Mapping and Data Reduction
2. Classification Algorithms
3. Ensemble Methods and Random Forests
4. Naive Bayes and Text Classification
5. Linear Regression
6. Clustering and Segmentation
7. Recommender Systems and Similarity Measures
8. Distributed Computing using MapReduce Algorithms

Learning outcomes:

1. Understand the purpose of data reduction and information extraction (e.g., associative statistics and data mapping).

2. Develop understanding and practical experience regarding reduction of massive data sets and data streams.

3. Understand the mechanics of distributed computing.

4. Ability to implement algorithms for processing massively large data sets. Ability to compute histograms, correlation matrices, and linear regression estimators using massively large data sets.

5. Understand the objectives of multiple regression and examining model assumptions. Ability to carry out and interpret hypothesis tests.

6. Competency using Python.

Textbook: None

Zoom:
Passcode: datasci
Meeting ID: 915 5613 1225

Getting Python:
You can download Python from Python.org. If you do not already have Python, I recommend instead installing the Anaconda distribution (www.anaconda.com), which already includes most of the tools that you need to do Machine Learning/Data Analysis

Homework and in-class projects:
Homework exercises emphasizing applications of the algorithms and/or theory, and in-class projects will be assigned bi-weekly.

Final Project:
In lieu of a final, students will be graded on their final project. Students must work on projects in groups of one, two or three individuals. Students are responsible for a written paper (75% of the project grade) and oral presentation (25% of the project grade).

Grading policy:

Your course grade will be based on homework, in-class projects, and the final project

Item	Percentage
Homework	40%
In-class projects	40%
Final project	20%

Student Conduct:

All students need to be familiar with the Student Conduct Code. You can find in the "A to Z Index" on the UM home page. All students must practice academic honesty. Academic misconduct is subject to an academic penalty by the course instructor and/or a disciplinary sanction by the University.

Accommodation:

The University of Montana assures equal access to instruction through collaboration between students with disabilities, instructors and Disability Services for Students (DSS). If you think that you may have a disability adversely affecting your academic performance, and you have not already registered with DSS, please contact DSS in Lommasson Center 154 or call 406.243.2243. I will work with you and DSS to provide an appropriate accommodation.

Safety:
1. Mask use is required within the classroom
2. Each student is provided with a cleaning kit. The expectation is that students will clean their personal workspace when they arrive for class, and before they leave the classroom.
3. Students are discouraged from congregating outside the classroom before and after class
4. Drinking liquids and eating food is discouraged within the classroom
5. Stay home if you feel sick and/or if exhibiting COVID-19 symptoms
6. If you are sick or displaying symptoms, please contact the Curry Health Center at (406) 243-4330

Digital Access:
Digital devices (like laptops and cell phones) are becoming increasingly important to success in college. I recognize that some students are unable to afford the cost of purchasing digital devices and that other students rely on older, more problem-prone devices that frequently break down or become unusable. I also recognize that those technology problems can be a significant source of stress for students. Given those challenges, I encourage students to contact me if you experience a technology-related problem that interferes with your work in this course.

Important note:
Announcements made in class are considered addenda to this syllabus. Make sure you stay informed as the progress of the class.