Math 460 Applied Data Mining. Fall 2022.

Instructor

Son Nguyen (snguyen4@bryant.edu)

Office Hours

You can find me at my office (Room A8 - Suite A) from 11:55AM-12:55PM on MWF. I am also available over Zoom (https://bryant.zoom.us/j/4419675207) during these times. If the office hours do not work, please feel free to schedule another time with me.

Course Website:

Canvas and https://bryantstats.github.io/math460/

Course Description

This course provides extensive hands-on experience with SAS Enterprise Miner software and covers the basic skills required to assemble analyses using the rich tool set of both packages. It also covers concepts fundamental to understanding and successfully applying data mining methods.

Prerequisites

MATH 350 or AM 332

Desired Learning Outcomes

After completing this course, you should be able to

- prepare data for analysis, including partitioning data and imputing missing values
- explain the algorithms of common predictive models such as decision trees, regression models, and neural networks.
- train, assess, and compare multiple predictive models such as decision trees, regression models, and neural networks.

Course Materials

The course will be based on the following texts. You are not required to buy these books. The reading materials will be provided on Blackboard.

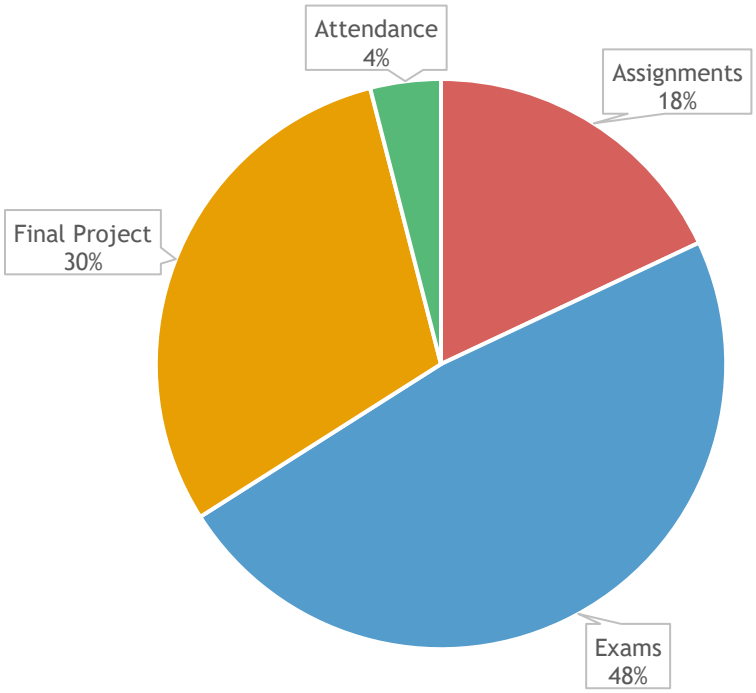
- Applied Analytics Using SAS® Enterprise Miner™ Course Notes. ISBN-13: 978-1612901398. ISBN-10: 1612901395

- Applied Predictive Models by Max Kuhn and Kjell Johnson. ISBN-13: 978-1461468486. ISBN-10: 1461468485

- Data Mining and Predictive Analytics (Wiley Series on Methods and Applications in Data Mining) 2nd Edition by Daniel T. Larose. ISBN-13: 978-1118116197. ISBN-10: 1118116194

- Introduction to Data Mining by by Pang-Ning Tan, Michael Steinbach, Vipin Kumar. ISBN-13: 978-0321321367. ISBN-10: 0321321367

Grades



Assignments

The assignments include, but not limited to, solving statistical problems, writing reports about a statistical technique, and presenting statistical analysis on a dataset.

Exams

The exams are written exams. Calculators and a formula sheet are allowed. There will be practice problems for each exam to help you prepare for the exam.

Final Projects

The final project offers you an opportunity to apply data mining techniques to an application. Projects can be done in teams of up to three students. If you have a project of such large scope and ambition that it cannot be done by a team of only three persons, you can propose doing a project in a team of four. The logistics and other details of the project will be updated on Blackboard.

Attendance

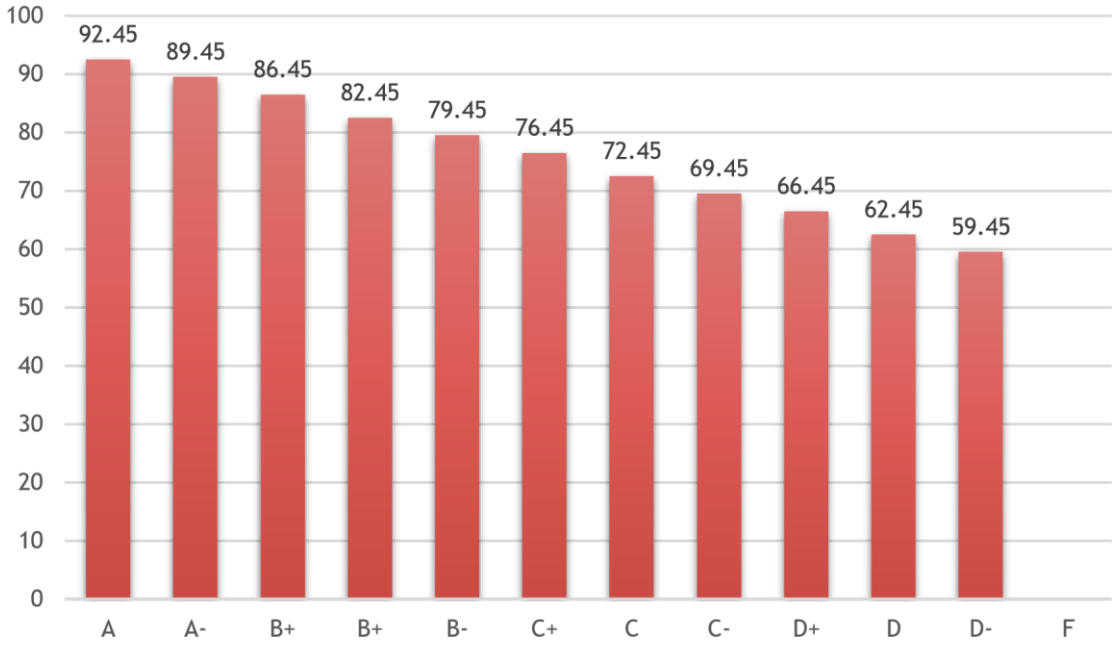
Attendance will be checked in every class period. Missing less than four classes will guarantee you the full credits for attendance.

Late Work

Late assignments are penalized 10% for each day late. You can resubmit your work to improve your score, but the late penalty will apply.

Grading Scale

A	92.45 - 100%	C	72.45 - 76.44%
A-	89.45 - 92.44%	C-	69.45 - 72.44%
B+	86.45 - 89.44%	D+	66.45 - 69.44%
B	82.45 - 86.44%	D	62.45 - 66.44%
B-	79.45 - 82.44%	D-	59.45 - 62.44%
C+	76.45 - 79.44%	F	Below 59.44%



Tentative Topic

- Overfitting and Model Tuning
- Classification Tree
- Random Forest
- Adaboost and Boosting Algorithm
- Gradient Boosting
- K-Nearest Neighbors
- Recommendation System
- ElasticNet
- Variable selection
- Neural Network and Deep Learning

Exam Tentative Schedule

Detail information about the exams will be posted on the course website. Follows are the tentative schedule.

Exam	Week	Date
Exam 1	4	09/26
Exam 2	8	10/24
Exam 3	12	11/21

Final Project

Final Project Presentation is scheduled at 10AM on December 19.

Academic Misconduct

Cheating will result in an "F" as your final grade and may result in your expulsion from the University. All cheatings will be reported to the Chair of the Mathematics Department and Academic Advising.

Regarding Diversity

In this course, and all your courses at Bryant, and throughout the Bryant learning community, we value and respect diversity. This includes differences in race, ethnicity, nationality, gender, gender identity, sexuality, socioeconomic status, ability, and religion.