MATH 795 Syllabus

Class and Instructor Details

MATH 795 - Data Science in Education - Fall 2021 (3 credits)

Instructor

Virgil U Pierce Email: virgil.pierce@unco.edu Cell: 956-249-0566

Office: Ross Hall 2239

Class Meetings

Tuesday and Thursday: 3:30p - 4:45p

Student Hours

Online and To be Determined.

Appointments can be made via: https://virgilpierce.youcanbook.me

Course Description

This course will introduce students to applications of data science specifically in the field of (mathematics) education. The aim of the course is two-fold: To teach students some of the tools and procedures for working with data science and data analytics in R specifically in the context of education research, and to train them in the topics so that they could teach an undergraduate introductory course in data science such as STAT 111.

Topics will include:

- Supervised Learning for Categorization and Regression
- Feature selection
- Visualizations
- Unsupervised Learning
- Structure and Design for Undergraduate Introductory Courses

Programing

We will be using R for this class. I highly recommend using R Studio which comes with some additional features to improve your workflow. Course notes will be prepared in R markdown. In industry, R markdown and Jupyter markdown for Python are becoming industry standards for communicating work involving programing and statistics and we need to be introducing undergraduate students to these tools.

You can install R Studio directly or via the Anacoonda package manger. The software is also available on the university Apporto Virtual Server.

- To install it on your own computer or laptop I recommend downloading Anaconda: https://www.anaconda.com/products/individual.
- You can log in to the Apporto Virtual Server with your UNC credentials.

Note that I have been teaching the undergraduate data science courses with Python. There is an active debate in the data science community about which software to use, though both are widely accepted standards. There are a ccouple of advantages to Python you should be aware of: Python is much bigger than just statistics/data science and it is the accepted computational tool for a wide variety of fields; and Python is available through Google Colab, which has become my suggested method for students to use it.

That said, both Python and R are good introductory programing langauges. Both of them provide for compact human readable code. R has an advantage when it comes to visualizations (unless you really know what you are doing). Both of them are too slow for some applications and for someone working seriously in industry learning a compiled software such as C++ will eventually be necessary.

Course Prerequisites

While this course will involve programing and statistics, I will not assume you have had a programming class before and will provide background material for any statistics results that we use. Data Science is an approach to statistics that emphasizes experimenting on collected data and as such provides a way to do similar computations one does in statistics but through experiments.

Textbook

Data Science References

Our primary text will be *Data Science in Education Using R* by Ryan A. Estrellado. Additional material for class will provided in R Markdown through Github; and finally we will refer to notes prepared for the STAT 411 class using Python: http://virgilpierce.org/datascience/frontmatter-1.html

R Reference

You may need an R reference. I recommend the O'Reiley texts, and there are a couple of good ones referring to R for Data Science. That said, while I like having a book to refer to for programming work, most of the time it does not have the detail I need and I end up using in-line help or on-line help anyway.

Learning Objectives / Outcomes for the Course

As a result of this course, students will expand their skill set to work and think with data. They will:

1. Be able to use machine learning algorithms for modeling data in categorical and regression problems.
2. Be able to present and visualize data.
3. Be able to evaluate models and discuss the process of identifying parameters and models for a problem.
4. Use resampling to revisit methods from a introductory statistics class.
5. Discuss the inclusion of Data Science in the undergraduate and high school curriculum.
6. Evaluate and discuss ehtical concerns with data science research generally, and specifically as it applies to research in education.

Course Outline

As a graduate level course, the outline will develop as class does, however we will start the class with an example of a machine learning model to motivate the rest of the topics.

Assessments

As with any mathematics class, doing the work is essentially to processing and learning the material. Furthermore the concepts covered in this class are not meant to be purely abstract things, but instead are actuall tools deployable in any human endeavor, even your hobies. You must do the activities in order to make progress.

https://www.youtube.com/watch?v=91G5SdW9tiQ

Homework

There will be some homework assignments asking you to explore topics from class, and expand on some of the work we do. These are intended to be done with a partner or two.

Also note that you do not necessarily need to have a perfect solution - for one thing there rarely is a perfect solution for this class, but rather you have spent some time thinking about how to get started.

Class Presentations and Flipgrid

You will give presentations to the class about the work we are doing.

We will use Flipgrid almost every class to share ideas and strategies on problems with each other.

Project

You will have a project you work on for the semester that can be any supervised learning problem. Details discussed in class. You will do this project with a team of three students. One of the main goals is I want to simmulate for you the team learning we do for STAT 411.

Project Videos

Each project will have an accompanying video where you explain your methods, results, and any issues that you have found.

Grading Policies

Your grade in the class will be based on homework, presentations and videos, and the project.

Grading and Feedback

I will be providing you feedback on the Homework, Class Activities and Presentations, the Projects and Project Videos. I am interested in seeing you make progreess as class continues. As such, you can revise and resubmit any assignment up until the announced deadline at the end of class.

Your final grade in class will be assessed primarily from the Projects and whether I see evidence that you have met the objectives for the class. Specifically: A grade of A will mean that you have met all six course objectives, and exceeded the standard on at least 4 of them; a grade of B will mean that you have met all six course objectives, and exceed the standard on at least 2 of them; a grade of C will mean that you have met all six course objects. Finally a grade of D will mean that at least one of the standards was not met, and a grade of F will mean that multiple standards were not met or that too many assignments were missing to allow me to assess a standard.

UNCO Policy Statements

Disability Resources

It is the policy and practice of the University of Northern Colorado to create inclusive learning environments. If there are aspects of the instruction or design of this course that present barriers to your inclusion or to an accurate assessment of your achievement such as time-limited exams, inaccessible web content, or use of videos without caption, please communicate this with your professor and contact Disability Support Services (DSS) at (970) 351-2289, Michener Library L-80 to request accommodations. Students can learn more about the accommodation process at https://www.unco.edu/disability-resource-center/accommodations/

Food Insecurity and Basic Needs

Research shows that college students experience food insecurity at higher rates than the American household rate, and that food insecurity can negatively impact academic performance and persistence. In recognition of this problem, UNC offers assistance to students facing food insecurity through an on- campus food pantry. The Bear Pantry is located in University Center 2166A, and is open for regular hours throughout the semester.

Please visit www.unco.edu/bear-pantry for more information.

Any student who faces challenges securing their food or housing and believes this may affect their performance in the course is also urged to contact Student Outreach and Support (SOS) for assistance. SOS can assist students during difficult circumstances which may include medical, mental health, personal or family crisis, illness or injury. SOS can be reached at sos@unco.edu or via phone at 970-351-2796.

Honor Code

All members of the University of Northern Colorado community are entrusted with the responsibility to uphold and promote five fundamental values: Honesty, Trust, Respect, Fairness, and Responsibility. These core elements foster an atmosphere, inside and outside of the classroom, which serves as a foundation and guides the UNC community's academic, professional, and personal growth. Endorsement of these core elements by students, faculty, staff, administration, and trustees strengthens the integrity and value of our academic climate.

UNC's Policies

UNC's policies and recommendations for academic misconduct will be followed. For additional information, please see the Student Code of Conduct at the Dean of Student's website http://www.unco.edu/dos/Conduct/codeofconduct.html. In the case of academic appeals, university procedures will be followed. For information on academic appeals, see https://www.unco.edu/trustees/pdf/board-policy-manual.pdf.