**Welcome to  the Fall, 2021 IST 707 Data Analytics at SU!**

Hi **@Class** ,

My name Dr. Gregory Block, I will be your instructor this semester. I am an adjunct faculty member at the iSchool. Here is a link to my profile:

https://ischool.syr.edu/gregory-block/

**First day of Asynchronous Coursework**: Wednesday, September 29th – Note, this is a flipped class. This means that you will prepare yourself prior to each class by taking advantage of the asynchronous material. Please complete Week 1 Coursework and HW before our first live session.

I have been told that there was a delay in your access to 2U, so I will understand if you need an extension on the first homework assignment.

**First Live Session**: Tuesday, October 12th, 6:00PM-7:30 PM EDT

**Class Outline and Due Dates**

1) HW is always due by 8:00 AM on the **Tuesday**of our **l**ive session for that week.

See our course schedule below. Note that we **will deviate from the schedule noted in the syllabus slightly**to better organize your time and provide you with extra time to work on your projects.

|  |  |  |  |
| --- | --- | --- | --- |
| Week | Topic | Textbook Readings | Submission items |
| 1 | Introduction to Data Mining | Ch.1 | HW1 |
| 2 | Data Exploration | Ch. 2-3 | HW2 |
| 3 | Association Rules | Ch. 6.1-6.3 | HW3 |
| 4 | Clustering | Ch. 8.1-8.3 | HW4 |
| 5 | Classification  algorithm: decision tree | Ch. 4.1-4.3 | HW5 |
| 6 | Model Evaluation | Ch. 4.4-4.6 | Practice Presentations |
| 7 | Classification algorithm: naïve  Bayes | Ch. 5.3 | HW6 |
| 8 | Classification algorithm: kNN, SVMs, random  forest | Ch. 5.2, 5.5 | HW7 |
| 9 | Text mining |  | ~~NO HW8 (due later!)~~ |
| 10 | Neural networks, part 1 |  |  |
| 11 | Neural networks, part 2 |  | HW8 Final project report (checkpoint 3) |

Note there are a few more changes to the course, which we will discuss during our first live session. For example, the async suggests you upload a video introduction. You can skip that - we will have a chance to meet each other during session breakouts.

**Class Book: Introduction to Data Mining: Tan, Steinbach, and Kumar (2005)**

Note: There are two editions of this book. This class uses the older edition, so please get that one (copyright 2005). I have both editions because this is a great book – very clear – excellent examples – and very applicable to Data Mining, Data Science, and Data Analytics.

In case your textbook is on order when class starts, I've uploaded a PDF version of chapter 1 to our Teams site, under the General channel, files tab, or [here](https://drive.google.com/file/d/1p15fuBs2eJgcvLoA4d53cw5pFn2su9tn/view?usp=sharing).

Apropos of Teams, I have created a Microsoft Teams site for our class. We will use 2U for the async material, submitting assignments and the gradebook, but for the live sessions we will use Teams. Here is the link to our Teams site.

<https://teams.microsoft.com/l/team/19%3aqxWkxV2eJzi8FVVhbtY1gGIIVeKypI5MNp-DKbQtZ3c1%40thread.tacv2/conversations?groupId=e76303c9-d3d3-43ec-978f-8e11b31d4b9d&tenantId=4278a402-1a9e-4eb9-8414-ffb55a5fcf1e>

Please let me know if you have any problems accessing the site. Sometimes the roster is not up-to-date when I add members to Teams, so I may have missed adding you!

This is a hands-on course with a heavy use of R Studio. I fully expect there will be issues and challenges when completing our lab assignments. If you have a question or issue, please post it to the Issues channel in our Teams site, so your fellow students can see the Q&A. If you have a personal or individual issue, then please email me directly, or reach me on Teams chat.

We have had a rough start to our term, so I won't be surprised if you do have challenges during the first week or so. Needless to say, I am happy to work with you to make sure our term gets back on track!

I look forward to meeting each of you on the 12th!

Dr. Block