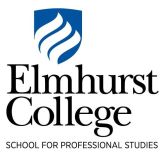
****

**MDS 556: Analytical Methods**

***Masters in Data Science***

**Fall Term 2016**

**Instructor and Contact Information**

James Kulich, Ph.D., Professor of Mathematics, Vice President and Chief Information Officer

Phone Number: (630) 617-3575

Email: [jimk@elmhurst.edu](mailto:jimk@elmhurst.edu)

The best way for you to contact me is by e-mail. I will respond to all e-mails submitted from 8 a.m. CT on Monday through 2 p.m. CT on Friday within 24 hours. I will respond to e-mails submitted at all other times within 48 hours. If you want to talk by phone or Adobe Connect, let me know by e-mail and I will work with you promptly to find a time that will work for both of us for a more in-depth conversation.

I will hold a weekly Webinar each Tuesday during out course from 6:45 to 7:45 p.m. CT. These Webinars are open to everyone in the course and will be fully available via Adobe Connect. Recordings of each week’s Webinar will also be available.

**Course Description and Culture**

MDS 556 is the second of two quantitative methods courses in the Elmhurst College Master of Data Science program. Our main goal for the course is to provide you with an enhanced set of mathematical tools that will allow you to build better and more robust models that work at any scale.

Each week will begin with a set of materials for you to explore that will include readings from selected texts and papers, illustrations of concepts to be considered, and when appropriate, case studies. Each Tuesday evening, at the scheduled weekly Webinar, I will present a lecture on the topic of the week that will both cover the week’s basic material and provide additional context. I will use applied case studies on occasion as well as taking deeper dives into the mathematics we will study. In order to best prepare for each Tuesday lecture, I encourage you to take a first pass through the week’s materials before our Tuesday meeting. The focus for the second half of each week will be on cementing your conceptual understanding of the material at hand as well as on gaining practice in applying the ideas to various datasets. You are encouraged to use all knowledge that you have developed thus far as well as any appropriate software tools.

We will use *Data Mining and Techniques* by Gordon S. Linoff and Michael J.A. Berry as our primary text for the course. I will also draw upon a variety of additional resources which will present our course material at several levels. Most readings will take an applied approach. However, especially for those among us with deeper interest in the mathematical theories, I will provide resources that offer an in-depth view.

You will have two options for your course project. In each case, the goal is to use new knowledge gained in this course to either develop or improve a predictive model. The option are:

* Extend the work you did in the MDS 549 Data Mining Project course
* Propose and complete a new predictive modeling project.

I will ask you to choose a path by the end of week three.

While I will lecture each week, my primary roles as your instructor will be those of coach and knowledge integrator. I will guide you in your efforts to pull the elements of your learning together in a way that makes sense for your own individual educational and professional goals. I fully expect that you as students will bring a range of experiences to the course. At any given time, anyone in our group might be the expert on a particular topic. I will work to create an environment in which sharing of your knowledge and expertise is safe, encouraged, and appreciated. I look forward to embarking on this next step of the data science journey with you!

**Course Outcomes**

By the end of this course, students are expected to:

* Evaluate strengths and shortcomings of various mathematical approaches to predictive modeling in various circumstances, especially when working at scale
* Demonstrate a good working knowledge of mathematical concepts covered in the course
* Identify opportunities in the modeling process to improve the quality of models by the use of more sophisticated techniques
* Design and execute a predictive model that illustrates good working knowledge of the course material;
* Communicate the results and implications of a complete analysis of complex data to both technical and non-technical audiences.

**Textbook and Learning Resources**

Our primary text is *Data Mining Techniques* by Gordon S. Linoff and Michael J.A. Berry, published by Wiley. We will also tap a variety or additional resources, many of which will be listed in the materials posted for each week. Materials at a variety of levels will be offered and used selectively to serve our learning goals. Some baseline resources that will be helpful include:

Blog posts, articles, and case studies from [Data Science Central](http://datasciencecentral.com). If you have not yet done so, please establish your free account.

Programming Collective Intelligence by Toby Segaran. It is available for purchase in hard copy or available at <http://4free-ebooks.com/ebook/programming-collective-intelligence/hmtsm23d6> for free download in .pdf format.

The Elements of Statistical Learning by Trevor Hastie, Robert Tibshirani, and Jerome Friedman. This book is made freely available in .pdf format by Stanford University at <http://statweb.stanford.edu/~tibs/ElemStatLearn/>. This is a deep and tough text. We will touch on it some but most of this will be reserved for those in our group who want to do a deeper dive.

Electronic Statistics Textbook by Statsoft. This free electronic text also covers the material at hand in considerable depth. It is available at <https://www.statsoft.com/Textbook>.

[Wikipedia](https://www.wikipedia.org/). Wikipedia is a surprisingly good resource for overviews of many concepts and contains a number of useful references in many cases. We will use some Wikipedia materials.

All of [Lynda.com](http://www.lynda.com/member.aspx) is now freely available to you. Go to the College Portal and click on the Lynda.com link under Applications. Your existing courses in Lynda.com should remain available to you together with anything else you might wish to explore

Tableau is available. The landing page is: <http://www/tableau.com/tft/activation>. The Desktop Key is **TD77-4EBC-A650-281B-A406**.You may, of course, use programming languages such as R and Python as well as any other software tools that may be available to you.

**Tentative Course Schedule**

|  |  |  |
| --- | --- | --- |
| **Module** | **Dates** | **Module Title** |
| **1** | **10/23 – 10/29** | **Getting Started: The Curse of Big Data, Correlation and Causation** |
| **2** | **1030 – 11/5** | **Essentials of Linear Algebra – With a Touch of Calculus** |
| **3** | **11/6 – 11/12** | **Clustering, Feature Selection, and a Little Monte Carlo**  **Project Proposal Due** |
| **4** | **11/13 - 11/19** | **Foundations of Multiple Regression** |
| **5** | **11//20 – 11/26** | **Modern Approaches to Regression** |
| **6** | **11/27 – 12/3** | **Time Series Analysis**  **Proctored Exam** |
| **7** | **12/4 – 12/10** | **Biological Methods: Neural Networks and Genetic Algorithms** |
| **8** | **12/11 – 12/17** | **Finishing: Analytical Techniques in Practice – Project Presentations** |

**Weekly Webinar and Problem Session Schedule**

The weekly Webinar will be on Tuesdays at 6:45 p.m. CT using the Adobe Connect platform. You must use the appropriate URL to enter into the virtual meeting room. You can find the URL within each module under the Weekly Webinar heading.

To join the webinar:

1. Click the **URL** contained inside the online classroom.
2. Select **Guest** and enter your full name.
3. Check your audio. Be sure your speaker volume is at an adequate level.
4. Use the interactive tools to actively participate.

**Grading Scale**

Elmhurst College uses the following grading scale.

A 4.0 A- 3.7 B+ 3.3 B 3.0 B- 2.7 C+ 2.3 C 2.0 C- 1.7 D+ 1.3 D 1.0 F 0.0

**Policy Regarding the Submission of Late Assignments**

All assignments are expected to be submitted by 11:59 p.m. CT Saturday of the week in which they are due. In the event that you are unable to complete work by the deadline, the instructor must be notified in advance.

**Statement Regarding Netiquette**

Netiquette refers to Online Etiquette, and it is the expectation of the College that all students will use appropriate online etiquette for interaction in online courses and using online communication tools, such as email and instant messaging. Common guidelines include:

* Avoid using strong or offensive language;
* Be forgiving. If something offends you, work with the instructor regarding the incident, rather than engaging in further enflamed discussion;
* Remember that your online messages and course postings live forever, so proof-read your work before you press Send.
* Avoid typing ALL IN CAPS. This is considered “shouting” online.
* Be concise; your points may be lost in a lengthy text;
* Give credit where credit is due: Use citations as appropriate;
* Avoid using Reply All in email messages.

**Disability Services**

Elmhurst College will make accommodations for students with disabilities based on the presentation of appropriate documentation. If you believe that you have a disability that may impact your work in this course, contact Dr. Corinne Smith, Disabilities Services Coordinator, at [corinne.smith@elmhurst.edu](mailto:corinne.smith@elmhurst.edu), 630-617-6448.

**Learning Center: Academic Support**

The Learning Center offers services to support the academic performance of all Elmhurst College students. Sessions are structured to promote principles of effective, self-regulated learning and academic management.  Areas of tutoring include math, writing, reading, study skills, and special test preparation (i.e., ACT, SAT, GRE, TAP). For more information, contact Mrs. Emmi McAdams, Math Specialist, at [emmim@elmhurst.edu](mailto:emmim@elmhurst.edu) , 630-617-5376, or Mrs. Susan Roach, Learning Center Director, at [susan.roach@elmhurst.edu](mailto:susan.roach@elmhurst.edu), 630-617-3155.

**Both Disability Services and the Learning Center are located in the Frick Center in Room 229.**

**Code of Academic Integrity**

Elmhurst College is a community of scholars.  Such a community places the highest value on academic honesty and integrity.  Therefore, Elmhurst College’s Code of Academic Integrity is designed to ensure:

1. A fair academic environment where students are evaluated on the scholarly merits of their work.
2. An academic environment where person integrity is fostered.
3. An understanding of standard academic practices when formulating ideas, conducting research, and documenting sources.

It is the responsibility of each student to know and understand this academic integrity policy.  There are general campus-wide definitions of academic honesty as well as department guidelines.  Students are responsible for educating themselves on this matter since sanctions for academic dishonesty may be severe.  Thus, Elmhurst College students in this course should act honestly in all academic work.  Please refer to the[***Elmhurst College E-Book***](http://issuu.com/elmhurstcollege/docs/ebook_2014-15_final_web/0) for additional information.

**Statement Regarding the Use of Plagiarism Detection Services**

Faculty members may at any time submit a student’s work for review by a plagiarism detection service.

**Non-Discrimination/Non-Harassment Policy**

The College believes that all employees and students should be able to work and learn in an educational environment free from discrimination and harassment.  Harassment includes unwelcome conduct, whether verbal, physical, written or graphic, that is based on protected group status such as race, color, national origin, religion, sex, sexual orientation, gender identity or expression, age, disability, genetic information, veteran’s status, or other protected status.  Please refer to the[***Elmhurst College E-Book***](http://issuu.com/elmhurstcollege/docs/ebook_2014-15_final_web/0) for additional information.

**Code of Conduct**

The College attempts to establish a climate which encourages the assumption of responsibility by everyone for their individual and collective actions with a minimum of rules and regulations.  It shall be the responsibility of each student to comply with appropriate standards of conduct and decorum befitting a member of an educational community.  Please refer to the[***Elmhurst College E-Book***](http://issuu.com/elmhurstcollege/docs/ebook_2014-15_final_web/0) for additional information.