University of Wisconsin-Milwaukee College of Health Sciences
Department of Health Informatics & Administration

HI 743
Predictive Analytics in Healthcare
Credits: 3

INSTRUCTOR

Ryan Gallagher, MA, Ad Hoc Professor
College of Health Sciences-Health Informatics and Administration
University of Wisconsin—Milwaukee

Telephone: (414) 251-8459
Email: gallag85@uwm.edu

Office Hours: By Appointment in Zoom

COURSE DESCRIPTION

With the recent rapid growth of electronic data in healthcare, it is now possible to apply techniques such as predictive analytics and machine learning to improve healthcare services. This includes building predictive models for prevention and diagnosis of diseases, to select the most effective treatments, to better utilize hospital resources, as well as to monitor patient data for early detection of problems. Besides patient data, these methods can also be used with public health data. This course will cover some of the most relevant topics in predictive analytics. We will also look at the applicability of such techniques in the healthcare context and where it could be applicable. We will examine the pros-and-cons and risks of using predictive modeling and machine learning in the healthcare context. For each of the machine learning methods, first theoretical concepts will be covered, followed by examples of their applications. This will be followed by a few detailed case studies of predictive analytics in healthcare. The course will include hands-on experience with a machine learning software. The course will provide an in-depth understanding of the machine learning methods and will also cover issues related to their evaluation and their adaption for medical applications.

COURSE OBJECTIVES

Machine learning and predictive analytics for healthcare is a very multi-disciplinary area that requires an understanding of statistics, basic computing skills, and an in-depth knowledge of the data domain or the context. There are different ways we can use machine learning techniques to build predictive data analytics models. In this course, students will:

1) Gain an understanding of which machine learning platform to use for different data contexts
2) Gain an overall understanding of the benefits and pitfalls of using machine learning and predictive modeling in healthcare
3) Gain an understanding of the different types of machine learning techniques
4) Gain an understanding of statistical learning

UWM credit hour policy:

This is a 3-credit course. You can expect to spend at least 9 hours/week in this course or a total of 144 hours over the semester. The time involved is roughly divided in thirds: reading and review of course materials, completion of assessments, and participating in course activities. The policy is available at the following link:
http://www4.uwm.edu/secu/docs/faculty/2838_Credit_Hour_Policy.pdf#zoom=75&page=2

COURSE MATERIAL:

Required Textbook: *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies* by John D. Kelleher, Brian Mac Namee and Aoife D' Arey, The MIT Press, 2015. Or the latest edition.

Supplementary Textbook: *Introduction to Statistical Learning with Applications in R (Second Edition)* by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, New York: Springer, 2013. Available here.

Other Required readings: These are from journal articles, research reports and monographs that will be posted to CANVAS. The readings may change as this is a very new and developing area and when I see new articles that are relevant, I will add them, so check CANVAS for each week's readings. All readings will be updated on Thursdays. The full text of all the readings will be available to download at the CANVAS content page.

PARTICIPATION/ATTENDANCE

Attendance requirement: Students will be expected to attend every class and read assigned readings.

LATE WORK POLICY

Late work for any assignment will not be accepted. The discussion boards are timed and will close at exact deadline time. Emailing the discussion post after the due date will not get any grades. Missing any of the deadlines will result in

zero points for that assignment. There is ample time for each assignment and there is also a lot of flexibility built in.

TECHNICAL REQUIREMENTS

Students are expected to access the CANVAS website for this course. Lecture materials, handouts, feedback on assignments, and grades will be posted at the site, so students are expected to check it regularly for course information.

If students have problems with access or navigation of the CANVAS site, they can contact the UWM Help Desk at 229- 4040, help@uwm.edu, or GetTechHelp@uwm.edu

We will be using 'R' for some hands-on projects/assignments in this class. You can download R from here: https://www.r-project.org/ and R studio from here: https://www.rstudio.com/ - This is an open source software and is available to download for free from the Internet. R also has extensive support on the web, so, you are encouraged to seek out whenever you need help with anything in R.

No additional software is required. A standard laptop computer with reliable internet access is sufficient. Access additional information for online & distance students at http://uwm.edu/onestop/online/

ASSIGNMENTS

- Participation/Attendance: 10%
- Assignments/Labs: 30%
- Interactive discussions/quiz (2): 30%
- Final project: 30%

Note: Pl. do not plagiarize – i.e. copy paste from other sources for any of these assignments. You are expected to write in your own words.

What is Plagiarism? How do you cite sources appropriately? Pl. read these http://www.plagiarism.org/plagiarism-101/what-is-plagiarism/
http://www.plagiarism.org/citing-sources/whats-a-citation/
https://owl.english.purdue.edu/owl/resource/589/2/
http://library.stevens.edu/plagiarism
http://writingcenter.utk.edu/for-students/citingsources/

Policies regarding plagiarism

- "I did not know this is plagiarism" is not a valid excuse – pl. read about plagiarism and if you need any clarifications, pl. contact me in the first week of class itself.
- Too much of the content in quotes is considered plagiarism. Avoid quotes as much as possible. Writing in your own words is the best policy.
- Do not copy and paste paragraphs and slightly modify the sentences so that you can cheat the plagiarism software. That is not how it works. I have been very successful in catching them.
- Once caught, you will be reported immediately for Academic misconduct.

Assignments & Quizzes

There will be a few assignments – these are to be decided. It can sometimes be a project or it can be a set of short answer questions in CANVAS.

Labs

During the second half of class, I will be leading a lab that shows application of the lectured topic. The expectation is that students code along with me in R and complete these labs. Completion for these will be graded.

Final Project

You will define your own final project. This is a team project – teams can be of 2 or 3 members each.

Gather a dataset and define a predictive analytics task in terms of what you will be predicting using which features. It can be something that is related to what you may be already working on, or will be doing in the future, for work, research, thesis, or for some other course. You are suggested to do something that is interesting to you, this way you will end up doing a better project. But avoid being too ambitious because that may lead to not finishing the project. More details will be available on CANVAS at the time of the final project.

Grading:

The project will be out of 30 points, and will have the following distribution:

5 points: Idea of the project (as described in the initial description)
5 points: What you did in the project
20 points: Project Report

EVALUATION/GRADING POLICY

Graduate Grade Scale

A: 93-100+ A-: 90-92.9 B+: 86-89.9 B: 83-85.9 B-: 80-82.9 C+: 76-79.9 C: 73-75.9
C-: 70-72.9 D+: 66-69.9 D: 63-65 D-: 60-62.9 F: Below 60

Extra credit

There is no additional extra credit for this course.

University Policy Statements

The Secretary of the University has a page dedicated to policies that includes information on policies for religious observances, incompletes, academic misconduct, grade appeal procedures, and final examinations, students called to military service, discriminatory conduct, and complaint procedures.
http://www4.uwm.edu/secu/news_events/upload/Syllabus-Links.pdf

Guidelines for Electronic & Wireless Devices in the Classroom

Electronic and wireless devices have become an integral part of modern life. These devices allow for ready access to knowledge and continuous connectivity. It is important to acknowledge not only the power of electronic/wireless devices but also the appropriate use of these devices. The link provided below explains the guidelines outlined that provide the College of Health Sciences community, including faculty, staff and students, with a framework directing the use of wireless and electronic devices in the classroom.
http://uwm.edu/healthsciences/students/guidelines-for-electronic-wireless-devices-in-the-classroom/

Honor Code

The Honor Code provides a framework for moral, ethical, and professional behavior for all members of the College of Health Sciences, including students, faculty, and staff. With all members of the College committed to upholding and promoting the tenets of the Honor Code, we will continue to work and learn in a supportive and stimulating environment. Commitment to this Honor Code supports the mission of the College of Health Sciences to prepare future health professionals, and conduct nationally recognized research in the health sciences.

Also available at http://uwm.edu/healthsciences/students/honor-code/

Syllabus is subject to change as this is a new area of study and so new topics might get added and old ones might get replaced – pl. keep checking the syllabus each week (on Thursdays).

Fundamentals of Machine Learning for Predictive Data Analytics = FMLPD

An Introduction to Statistical Learning with Applications in R = ISLR

COURSE SCHEDULE

WEEK / DATE	TOPICS	READINGS
Week #1 January 23	Introduction to HI 743 - Predictive Analytics in Healthcare	
Week #2 January 30	Machine Learning for Predictive Analytics	Chapter 1 (FMLPD) Chapter 2 (ISLR) Readings & materials in CANVAS
Week #3 February 6	Intro to R language	Chapter 2.3 (ISLR) Assignment 1
Week #4 February 13	Data Exploration	Chapter 3 (FMLPD) Readings & materials in CANVAS
Week #5 February 20	Error-based learning Regression – linear	Chapter 7 (FMLPD) Chapter 3 (ISLR) Readings & materials in CANVAS
Week #6 February 27	Regression - logistic	Chapter 7 (FMLDP) Chapter 4 (ISLR) Readings & materials in CANVAS
Week #7 March 6	Case Study	Readings & materials in CANVAS
Week #8 March 13	Information based learning	Chapter 4 (FMLDP) Chapter 8 (ISLR) Readings & materials in CANVAS
Week #9 March 20	**SPRING BREAK (March 16th – March 23rd)**	

Week #10 March 27	Similarity-based learning	Chapter 5 (FMLDP) Chapter 12 (ISLR) Readings & materials in CANVAS Assignment 3
Week #11 April 3	Introduction to Probability theory Probability- based learning	Chapter 6 (FMLPD) Chapter 4 (ISLR) Readings & materials in CANVAS Assignment 4
Week # 12 April 10	Neural Networks and Deep Learning	Chapter 10 (ISLR) Readings & materials in CANVAS
Week # 13 April 17	Evaluation	Chapter 8 (FMLPD) Chapter 5 (ISLR) Readings & materials in CANVAS
Week # 14 April 24	The Art of Machine Learning for Predictive Data Analytics	Chapter 11 (FMLPD) Readings & materials in CANVAS
Week # 15 May 1	Work on your final project	
May 9	Final project papers due	