MSCI:6070 Summer 2019

Data Science

INSTRUCTOR Nick Street
nick-street@uiowa.edu, (319) 335-1016
S358 PBB

GRADER TBD

CLASS MEETING TIMES 6:00-9:40 W
May 15 – July 24

OFFICE HOURS Before and after class or by appointment

COURSE SITE To access the course site, log into Iowa Courses Online (ICON) using your Hawk ID and password.

Program Goals

The administrative home of this course is the Tippie MSBA program, which governs academic matters relating to the course. The Tippie MSBA has learning goals that drive decisions about curriculum and assignments within courses. These goals are that graduates will: 1) exhibit knowledge and skills relevant to data and its application in business; 2) create and communicate solutions to data-related business problems that impact their organizations and communities; 3) understand and contemplate ethical and privacy issues arising in their own work; 4) demonstrate the ability to be effective team members in a diverse and complex world.

Course Description and Goals

In this course you will learn the basics of data science, particularly data mining, as applied to business problems. Our focus will be on the process of turning raw data into intelligent decisions, and on the algorithms that are commonly used to build predictive models and find relevant patterns in data. Class sessions will be a mixture of conceptual material and hands-on practice.

Learning objectives:

1. To understand the knowledge discovery process and how to apply it.
2. To master the details of various supervised and unsupervised machine learning algorithms.
3. To understand the underlying principles of predictive modeling such as complexity control.
4. To learn appropriate methods for fairly evaluating the quality of constructed models.
5. To gain practical experience in using a software package to preprocess data, build and evaluate models, and interpret the results.

Media / System Requirements

Each student will be expected to bring a laptop computer to each class meeting. Any modern operating systems should be fine, although I will be using Windows for examples. The following software will need to be installed:

- R: http://www.r-project.org/
- R Studio: http://www.rstudio.com/
- Rattle: http://rattle.togaware.com/

Textbook / Materials

The required textbook(s) for this course are:

Data Science for Business: What you need to know about data mining and data-analytic thinking, F. Provost and T. Fawcett. O'Reilly, 2013.

Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery, G. Williams. Springer, 2011.

Readings from other sources may be made available in class.

Grading Criteria

50% Homeworks
45% Exams
5% Class participation

The score - grade mapping will be determined by the instructor at the end of the course. The usual 90/80/70 cutoffs will be used as a baseline, but may be adjusted.

Course Work

Homework assignments: You will be assigned 4-5 application homework problems. Each assignment will require you to analyze a supplied dataset using Rattle. You will be expected to correctly apply the concepts that have been presented in lecture, thoroughly analyze the results of the computational models, and write up your conclusions in a professional manner.

Exams: There will be a comprehensive final exam given at the end of the class. There may or may not be a midterm exam, depending on the collective opinion of the students.

Course Calendar

The following calendar is approximate and may be modified as the course proceeds. Reading assignments are shown on the day when the corresponding material is expected to be presented; homeworks are shown on the day they are expected to be assigned.

DATE	*TOPICS*	*READINGS AND ASSIGNMENTS*
Week 1	KD process; data preprocessing	DSB ch. 1, 2 DMRR ch. 1-5, appendix A
Week 2	Preprocessing, cont.; Classification; decision trees	DSB ch. 3 DMRR ch. 8, 11 Homework 1
Week 3	Model evaluation; logistic regression	DSB ch. 7 (skim expected value framework) DMRR ch. 7, 15.1-15.3 Homework 2

Week 4	Support vector machines; artificial neural networks	DSB ch. 4-5 DMRR ch. 14 Supplemental readings
Week 5	ANNs, continued; ROC, lift and profit curves	DSB ch. 8 Homework 3
Week 6	Midterm exam (?); bias/variance tradeoff; ensemble methods	DMRR ch. 12-13
Week 7	Distance measures; k-means clustering	DSB ch. 6 DMRR ch. 9
Week 8	Hierarchical clustering; cluster interpretation	Homework 4
Week 9	Association analysis	DSB ch. 12 (first sections) DMRR ch. 10 Homework 5
Week 10	Special topics TBD; exam review	TBD
Week 11	Final Exam	none

Course-Specific Policies and Guidelines

Attendance: Attendance is not required but is strongly encouraged. Class attendance is essential for learning the material. Class material will be provided via lecture capture and smart board notes. You will be responsible for knowing everything discussed in class, and everything that is posted to the ICON site.

Assignments: Students are allowed to collaborate freely on homework assignments, but you should write the final submission independently to make sure you understand each step. Identical homework submissions or exam answers will be considered evidence of academic dishonesty.

Late assignments: All assignments are expected to be turned in at the beginning of class on their assigned due date. Late assignments will be penalized 10% of the available points, and another 10% will be deducted for every 24-hour period after the original due date.

Exams: Students are expected to take the exams at the regularly scheduled times, unless permission has been granted by the instructor at least 2 weeks in advance in writing. Exams are individual work and collaboration on exams is not permitted under

any circumstance. All exams are paper-based and closed book. Use of any materials during the exam is prohibited and will be grounds for a grade of zero on the exam.

Collegiate Policies and Guidelines

The administrative home of this course is the Tippie College of Business, which governs academic matters relating to the course such as the add/drop deadlines, the second-grade-only option, issues concerning academic misconduct, and how credits are applied for various graduation requirements. Different colleges might have different policies. As a registered student in a course in the Tippie College of Business, you are responsible for the collegiate policies posted below.

Please note that students of this program are subject to the policies, procedures, and requirements of the Graduate College and The University of Iowa as well. For additional details, students should consult: the Manual of Rules and Regulations of the Graduate College (http://www.grad.uiowa.edu/graduate-college-manual), applicable sections of The University of Iowa Catalog (http://www.registrar.uiowa.edu/registrar/catalog/) and Policies and Regulations Affecting Students (http://dos.uiowa.edu/policy-list/current/).

Tippie Honor Code: The Tippie College of Business has an Honor Code, and you must abide by it in completion of all assignments. Integrity is a reflection of your character and is critical for creating meaningful and lasting relationships. One part of integrity is abstaining from acts like cheating, so cheating on any assignment in this class will result in an appropriate consequence, usually a zero for the grade in question and, if that penalty does not reduce the grade, a penalty of a full letter grade reduction. In addition, all incidents of cheating will be reported to the appropriate academic offices, and the student may be placed on disciplinary probation, be suspended, or even permanently expelled, depending on the severity of the offense.

If a student has been found in violation of this policy, they will first be notified directly, then I will report to the appropriate program office. Faculty and students can report Academic Misconduct via the college website.

Privacy: As part of this course, your name will be revealed to other students in the class. If you have concerns, please contact the instructor immediately.

Lecture Capture Policy: As a service to students, the program offers Lecture Capture of most of its courses. This allows students to access video and audio recordings of the class, after the class has been recorded. As a student, you can opt out of being visible in the recording. Unless otherwise noted, this course is being recorded, and will be available to all enrolled members of the class under a password protected website (the ICON course site). The instructor may choose at various points (i.e. during exams or guest speakers) to not record the class.

Fairness and Freedom of Expression: Every student is entitled to the same intellectual freedom I have. I will respect that freedom, and I am obliged to protect your freedom to learn, regardless of your religion, race, sex, sexual orientation, gender identity, or political views, or on your agreement or disagreement with my positions pertaining to matters of controversy within the discipline. I will do my best to provide you with a fair and impartial evaluation of your work, consistent with articulated standards for this course.

Complaint Procedures: If at any time you have concerns about this class or your performance in it, please contact me directly. If you do not feel that your concern has been resolved satisfactorily, you should contact the Department Executive Officer (DEO) who oversees the department offering this course (Barry Thomas, barrett-thomas@uiowa.edu). If that does not resolve the complaint, you may contact the Assistant Dean of Specialty Master's Programs, Dave Deyak, dave-deyak@uiowa.edu.

University Policies and Guidelines

As a registered student in a course through The University of Iowa, the following University policies apply to you.

Accommodations for Disabilities: The University is committed to provide an educational experience that is accessible to all students. If you have a diagnosed disability or any other condition that would impair your ability to complete the course requirements as stated above, please inform me as early in the semester as possible, but no later than two weeks prior to the scheduled activity. Students needing accommodations must register with Student Disability Services, (SDS): https://sds.studentlife.uiowa.edu/students/apply to obtain a Student Academic Accommodation Request (SAAR) form. The form will specify what course accommodations are reasonable for that student. The office is located at 3100 Burge Hall, (319) 335.1462.

Mental Health: Students are encouraged to be mindful of their mental health and seek help if they are feeling overwhelmed or incapable of meeting course expectations. For assistance with the class, students are encouraged to talk to the faculty member. For additional advice or support, students are encouraged to contact University Counseling Services at University Capital Centre Suite 1950 or 3223 Westlawn South (319-335-7294) for same day appointments (Mon-Fri. 8AM-4:30PM). After hours, we encourage you to call the Johnson County Crisis Line at 319.351.0140 or 911 if you are in immediate danger.

Sexual Harassment: Sexual harassment subverts the mission of the University and threatens the well-being of students, faculty, and staff. The University will not tolerate sexual harassment, nor will it tolerate unwelcomed behavior of a sexual nature toward

members of the University community when that behavior creates an intimidating or hostile environment for employment, education, on-campus living, or participation in a University activity. As a member of the university community you have a responsibility to report concerns of sexual harassment immediately at the Office of the Sexual Misconduct Response Coordinator: https://osmrc.uiowa.edu/.

.

Multicultural Holidays: Students compelled by their religious convictions to refrain from attending class on specific days must request instructor permission during the first few days of the session. If this request is denied, a student may address concerns according to University Procedures.

Sustainability: The University is committed to demonstrating sustainability practices within all facets of the institution. Students are encouraged to utilize recycled materials and use as few natural resources as possible (such as by making use of digital forms of note taking, as appropriate). Recycling of all paper materials is expected. Learn more at the Office of Sustainability.