

## **Data Science**

**INSTRUCTOR** Jeffrey Ohlmann, Ph.D., Associate Professor  
[jeffrey-ohlmann@uiowa.edu](mailto:jeffrey-ohlmann@uiowa.edu), (319) 335-0837

**CLASS** 6:00pm – 9:40pm, Wednesdays  
**MEETING** May 15 – July 24  
**TIMES** 111 BLC Birchwood Fields Learning Center, Davenport

**OFFICE** Before and after class or by appointment  
**HOURS**

**COURSE** To access the course site, log into [Iowa Courses Online \(ICON\)](#) using your  
**SITE** 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 in order to participate in class activities. Any modern operating system should be fine, although course instruction will be based on a PC-based Windows machine. 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/>

If you need assistive technologies that has different computer and technology requirements, then you will be responsible for making those arrangements. Please check with [Student Disability Services](#) to determine the requirements for specific technologies and for assistance if necessary.

## 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*
  - ISBN: 978-1449361327
  - Author: F. Provost and T. Fawcett
  - Publisher: O'Reilly
  - Copyright Year: 2013
- *Data Mining with Rattle and R: The art of excavating data for knowledge discovery*
  - ISBN: 978-1441998897

- Author: G. Williams
- Publisher: Springer
- Copyright Year: 2011

## Grading Criteria

Final course grades will be assessed based on the student's performance in the following activities:

| Graded Item   | % of Final Grade |
|---------------|------------------|
| Homework Sets | 24%              |
| Participation | 6%               |
| Exams         | 70%              |

The Tippie College of Business guidelines suggest that an appropriate grade distribution would consist of less than 50% of the students in the A range, greater than 50% of the students in the B range, and grades in the C, D, or F range as merited. Note that the exact grade distribution will depend on the class's performance and plus / minus grading will be used. The instructor reserves the right to adjust the grade distribution as appropriate.

If you have a concern regarding the grading of an assignment, you must submit your re-grade request in writing attached to the assignment within one week of its receipt. The instructor reserves the right to re-grade the entire document.

## Course Work

### Homework Sets:

This course includes a series of homework assignments to be completed by the individual student. Students may discuss and collaborate, but each student should create and write their own solution independently to ensure understanding. Each assignment will involve a thorough analysis of a supplied data set using course concepts and will require a lucid description of derived conclusions. Assignments are due in the appropriate ICON Dropbox by the date and time listed on the assignment. As solutions to homework sets are generally posted shortly after the deadline, no late submissions are accepted.

### Participation:

Attendance at all class meetings is expected and no make-up credit for absences is provided. Students are expected to give adequate advance notice of an excused absence, make up the missed work and make every effort to avoid a class conflict. To

ensure legitimacy, students may be asked to provide documentation upon request.

### Exams:

This course includes two computer-based exams that will require you to use the software to solve problems conceptually similar to those discussed in class. Students will have an entire class session to work on the exam. Students may use course materials while working on the exam, but will not be allowed to access the internet. 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.

## Course Calendar

| Date    | Lecture Topic*  | Reading   | Assignment | Due Date                      |
|---------|---|---|------------|-------------------------------|
| May 15  | Knowledge Discovery Process;<br>CRISP-DM;<br>Intro to Rattle;<br>Data Exploration & Visualization | DSB Ch. 1, Ch. 2;<br>DMRR Ch. 1-5,<br>Appendix A                |            |                               |
| May 22  | Similarity Measures;<br>k-Means Clustering  | DSB Ch. 6   |            |                               |
| May 29  | Hierarchical Clustering   | DMRR Ch. 9  | Set A      | 6:00pm on 5/29<br>(Wednesday) |
| June 5  | Association Analysis  | DSB Ch. 12<br>(first two sections);<br>DMRR Ch. 10              |            |                               |
| June 12 | <b>Exam I</b>   |   | Set B      | 11:59pm on 6/10<br>(Monday)   |
| June 19 | Classification;<br>Classifier Performance Evaluation;<br>Decision Trees                           | DSB Ch. 3, Ch. 7<br>(first two sections);<br>DMRR Ch. 8, Ch. 11 |            |                               |
| June 26 | Linear Model-Based Classifiers  | DSB Ch. 4;  |            |                               |
| July 3  | Support Vector Machines;<br>Artificial Neural Networks  | DMRR Ch. 14   | Set C      | 6:00pm on 7/4<br>(Wednesday)  |
| July 10 | Overfitting Avoidance;<br>Curves: ROC, Lift, Profit   | DSB Ch. 5, Ch. 8;<br>DMRR Ch. 15                                |            |                               |
| July 17 | Ensemble Methods  | DMRR Ch. 12-13  | Set D      | 6:00pm on 7/17<br>(Wednesday) |
| July 24 | <b>Exam II</b>  |   |            |                               |

---

\*Course schedule is tentative and subject to change.

## 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](https://sds.studentlife.uiowa.edu/students/apply), (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 \(319-335-7294\)](https://www.uiowa.edu/counseling) 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/](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](#).