

WAYNE STATE UNIVERSITY

COLLEGE OF ENGINEERING

DSA 6000: Data Science and Analytics - 3 Credits Course Syllabus – Fall 2020

- Instructor:** Yanchao Liu, Ph.D.
Assistant Professor, Department of Industrial & Systems Engineering
yanchaoliu@wayne.edu
- Teaching Assistant** N/A
- Time & Location:** Tuesdays and Thursdays 5:30 – 6:45 pm, Zoom meetings (link posted on Canvas)
- Office Hours:** By Appointment (Chat online using Microsoft Teams or Zoom)
- Web Sites:** <http://canvas.wayne.edu>
- Description:** A fundamental course covering basic data science and analytics concepts through case studies, success stories, and a semester project that cuts across all course modules. Students will be exposed to a variety of techniques for data exploration and modeling. Open to Data Science and Business Analytics majors only.
- Learning Outcomes:** After completing the course, students will:
- Know the elements of a data science and analytics project life-cycle, from business need to solution deployment
 - Understand core data science and analytics principles, methodology and tools
 - Understand some main applications of data science
 - Evaluate predictive models: how well does a model predict a phenomenon
 - Achieve an intermediate level of using R for data analysis, visualization and programming
 - Apply analytic thinking and expected value framework to business problems
- Prerequisites:** College-level linear algebra, calculus and statistics. Familiarity with a programming language.
- Textbook:** James, Witten, Hastie and Tibshirani, *An Introduction to Statistical Learning with Applications in R (ISLR)*. [[Free PDF](#)] or buy a hard copy at WSU Bookstore.
Provost and Fawcett, *Data Science for Business (DSFB) - What You Need to Know about Data Mining and Data-Analytic Thinking*, O'REILLY, 2013
- References:** Bishop, C. *Pattern Recognition and Machine Learning*. Springer, NY: New York, 2007 (ISBN: 0387310738). [[Free PDF](#)].
- Additional tutorials and journal papers will be distributed in the class as needed to complement the material from the textbooks in the areas of Deep Learning, Decision Trees, Random Forests, Data Mining, and others.
- Software:** RStudio, Microsoft Excel, Tableau Desktop, Python, Bash
- Grading:**
- | | |
|-----------------------------------------------------------------------------|----------------------|
| Homework | 45% |
| Mid-term Exam | 15% |
| Final Exam | 15% |
| Course Project | 20% |
| Participation | 5% |
| [Bonus: Short Bio & Picture | 0.5% if sent by 9/7] |
| Final Grade: A (>= 93), A- (>=89), B+ (>=85), B (>=80), B- (>=70), C (< 70) | |
- Individual projects, exams, and special assignments might be curved and changed with regard to importance (i.e., in points), at the discretion of the instructor. Project reports and assignment reports have to be typed, and when feasible, results have to be justified and thoroughly summarized (without appending lots of pages of output). Reports have to be submitted at the beginning of the class on the due date. Late reports will receive lower grades.

Homework Policy:	Homework assignments will be posted on course website and announced in class. Homework will be due at the beginning of the class period. Homework submitted late by one class will be evaluated at 50% credit and late by two (or more) classes will not receive any credit. Homework should be submitted on clean sheets and will be evaluated based on completeness. Depending on the size of the class and length of the homework, individual homework grades could be based on randomly sampled problems. When feasible, solutions for homework problems will be posted on the course website. When appropriate, students are encouraged to use Excel templates (available from Blackboard website) to verify calculations and results. In case the student is unable to attend class for submitting the homework, e-mail the work to the Teaching Assistant.
	<u>On Working Together:</u> Collaboration among students during the conceptualization and formulation process of homework assignments is allowed. However, all final homework assignments must be completed and written up individually. Violation of this policy will certainly lead to a failing grade for the homework and more stringent action is likely to be taken (e.g., assigning an "F" grade for the course or even dismissal from the MS program).
Exams:	Exams will focus on the material in the chapter readings, case studies and lectures. They will be open-book open-computer and will contain multiple choice, short answer questions and mini-case analyses. Makeup examinations must be requested in writing and must be the direct result of a medical or work-related issue.
Semester Project:	This is a team project with up to five students per team. The project has three phases: 1. Topic identification and proposal; 2. Mid-term review and presentation; 3. Final report and team presentation. More guidance, format template and feedback will be given to teams as the semester proceeds. All members of a team will receive the same project scores.
Attendance Policy:	Students attending any given class are required to join the class within the first five minutes to minimize any class disruptions.
Online Instruction due to COVID-19	Class meetings will be held using Zoom via Canvas. Class material (including PowerPoint slides, R code and problem descriptions) will be posted on Canvas before class. The class will proceed in the usual way: I will go over the presentation slides and R code, with frequent pauses for questions. There will be timed quizzes during class time to assess the learning outcomes and to track class attendance. There will be group break-out sessions for team project discussions. Project mid-term report and final presentation will be in the form of team presentations online.
Religious Holidays:	Because of the extraordinary variety of religious affiliations of the University student body and staff, the Academic Calendar makes no provisions for religious holidays. However, it is University policy to respect the faith and religious obligations of the individual. Students with classes or examinations that conflict with their religious observances are expected to notify their instructors well in advance so that mutually agreeable alternatives may be worked out.
Student Services:	<ul style="list-style-type: none"> ○ <i>The Academic Success Center</i> (1600 Undergraduate Library) assists students with content in select courses and in strengthening study skills. Visit http://success.wayne.edu for schedules and information on study skills workshops, tutoring and supplemental instruction (primarily in 1000 and 2000 level courses). ○ <i>The Writing Center</i> is located on the 2nd floor of the Undergraduate Library and provides individual tutoring consultations free of charge. Visit http://clasweb.clas.wayne.edu/writing to obtain information on tutors, appointments, and the type of help they can provide.
Class Recordings:	Students need prior written permission from the instructor before recording any portion of this class. If permission is granted, the audio and/or video recording is to be used only for the student's personal instructional use. Such recordings are not intended for a wider public audience, such as postings to the internet or sharing with others. Students registered with Student Disabilities Services (SDS) who wish to record class materials must present their specific accommodation to the instructor, who will subsequently comply with the request unless there is some specific reason why s/he cannot, such as discussion of confidential or protected information.
Academic Dishonesty – Plagiarism and Cheating:	Academic misbehavior means any activity that tends to compromise the academic integrity of the institution or subvert the education process. All forms of academic misbehavior are prohibited at Wayne State University, as outlined in the Student Code of Conduct (http://www.doso.wayne.edu/student-conduct-services.html). Students who commit or assist in committing dishonest acts are subject to downgrading (to a failing grade for the test, paper, or other course-related activity in question, or for the entire course) and/or additional sanctions as described in the Student Code of Conduct.

- **Cheating**: Intentionally using or attempting to use, or intentionally providing or attempting to provide, unauthorized materials, information or assistance in any academic exercise. Examples include: (a) copying from another student's test paper; (b) allowing another student to copy from a test paper; (c) using unauthorized material such as a "cheat sheet" during an exam.
- **Fabrication**: Intentional and unauthorized falsification of any information or citation. Examples include: (a) citation of information not taken from the source indicated; (b) listing sources in a bibliography not used in a research paper.
- **Plagiarism**: To take and use another's words or ideas as one's own. Examples include: (a) failure to use appropriate referencing when using the words or ideas of other persons; (b) altering the language, paraphrasing, omitting, rearranging, or forming new combinations of words in an attempt to make the thoughts of another appear as your own.
- **Other** forms of academic misbehavior include, but are not limited to: (a) unauthorized use of resources, or any attempt to limit another student's access to educational resources, or any attempt to alter equipment so as to lead to an incorrect answer for subsequent users; (b) enlisting the assistance of a substitute in the taking of examinations; (c) violating course rules as defined in the course syllabus or other written information provided to the student; (d) selling, buying or stealing all or part of an un-administered test or answers to the test; (e) changing or altering a grade on a test or other academic grade records.

Student Disability Services:

If you have a documented disability that requires accommodations, you will need to register with Student Disability Services for coordination of your academic accommodations. The Student Disability Services (SDS) office is located at 1600 David Adamany Undergraduate Library in the Student Academic Success Services department. The SDS telephone number is 313-577-1851 or 313-202-4216 for videophone use. Once you have met with your disability specialist, I will be glad to meet with you privately during my office hours to discuss your accommodations. Student Disability Services' mission is to assist the university in creating an accessible community where students with disabilities have an equal opportunity to fully participate in their educational experience at Wayne State University. You can learn more about the disability office at <http://studentdisability.wayne.edu/>. Students who are registered with Student Disability Services and who are eligible for alternate testing accommodations such as extended test time and/or a distraction-reduced environment should present the required test permit to the professor at least one week in advance of the exam. Federal law requires that a student registered with SDS is entitled to the reasonable accommodations specified in the student's accommodation letter, which might include allowing the student to take the final exam on a day different than the rest of the class.

Course Drops and Withdrawals:

In the first two weeks of the (full) term, students can drop this class and receive 100% tuition and course fee cancellation. After the end of the second week there is no tuition or fee cancellation. Students who wish to withdraw from the class can initiate a withdrawal request on Pipeline. You will receive a transcript notation of WP (passing), WF (failing), or WN (no graded work) at the time of withdrawal. No withdrawals can be initiated after the end of the tenth week. Students enrolled in the 10th week and beyond will receive a grade. Because withdrawing from courses may have negative academic and financial consequences, students considering course withdrawal should make sure they fully understand all the consequences before taking this step. More information on this can be found at: <http://reg.wayne.edu/pdf-policies/students.pdf>

Deferred Grade:

A grade of 'I' can only be assigned if all of the following criteria are met:

1. the student IS NOT currently failing the class and,
2. there is NOT a substantial quantity of work yet to be completed,
3. there is no extra work required of the instructor beyond the normal duties of grading the paper/exam,
4. there is no need for the student to attend the class in subsequent terms.

The final decision to assign an incomplete grade rests with the instructor. An 'I' grade MUST be made up within one year of assignment of the grade.

Tentative Schedule

We plan to cover ISLR chapters 2,3,4 and 8 with hands-on exercises in R, DSFB chapters 1-4, 7, 9, 11-13, the basic usage of Tableau (tutorial from Tableau website) and the basics of mathematical programming using GAMS. Lecture and lab materials including slides and computer codes will be uploaded on Canvas. However, I highly recommend that you obtain a copy (PDF or hardcopy) of both ISLR and DSFB and read the corresponding chapters as we go along.

Date	Topic	Reading / Assignment
9/1/2020	Course overview	Read stats review post, Read DSFB Ch 1
9/3/2020	Get started with R programming	Install RStudio, follow R tutorial
9/8/2020	Review statistical concepts	Form project groups, Read ISLR 2.1 – 2.2
9/10/2020	Exercise in R: data analysis and visualization	Read DSFB Ch 2, HW 1 due
9/15/2020	Expected value framework for analytical engineering	Read DSFB Ch 11
9/17/2020	Multiple linear regression	Read ISLR 2.3
9/22/2020	Other data science tasks and techniques	Read ISLR 3.1, DSFB Ch 12, HW 2 due
9/24/2020	Linear model diagnostics	Read ISLR 3.6.1 – 3.6.3
9/29/2020	Case study: regression / R Markdown	Read ISLR 3.3 – 3.6
10/1/2020	Groups present project proposal / discussion	Read DSFB Ch 3, HW 3 due
10/6/2020	Classification methods, logistic regression	Read DSFB Ch 4
10/8/2020	Logistic regression, confusion matrix, ROC curve	Read 4.1 – 4.3
10/13/2020	Naïve Bayes, LDA, QDA	Read DSFB Ch 9, HW 4 due
10/15/2020	Case study: classification, model selection and validation	
10/20/2020	Midterm Exam	
10/22/2020	Groups present project progress / discussion	
10/27/2020	Decision trees, tree induction	Read ISLR 8.1
10/29/2020	Tree ensembles, bagging, random forests, boosting	Read ISLR 8.2 – 8.3, HW 5 due
11/3/2020	Exercise using GUI tools in R	
11/5/2020	Visualize data using Tableau	Install Tableau, watch tutorial
11/10/2020	Data Science and Business Strategy	Read DSFB Ch 13, HW 6 due
11/12/2020		
11/17/2020	Turning insights into decisions: decision science overview	Read handout, install GAMS
11/19/2020	Basics of mathematical programming	
11/24/2020	Case study: mathematical optimization	HW 7 due
11/26/2020	Holiday, no class	
12/1/2020	Review session	Presentation slides due
12/3/2020	Project presentation	
12/8/2020	Project presentation	Project reports due
12/10/2020	Final Exam	