WAYNE STATE UNIVERSITY
COLLEGE OF ENGINEERING

DSA 6000: Data Science and Analytics - 3 Credits
(Data Mining, Business Analytics)
Course Syllabus - Fall 2017

Instructor: Yanchao Liu, Ph.D.
Assistant Professor, Industrial & Systems Engineering Department
??@wayne.edu | Tel: TBA

Time & Location: Tuesday & Thursday 5:30 – 6:45 pm, Room 0120, Manoogian Hall (MANO), 906 W Warren Ave

Office Hours: By Appointment, Room ???, 4815 Fourth Street (MEB), Wayne State University, Detroit, MI 48202

Web Sites[1]: http://blackboard.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 and software tools for data exploration and modeling. Open to Data Science and Business Analytics majors only.

Course Learning Outcomes: After completing the course, students should be able to:

- Discuss the elements of a data science and analytics project life-cycle, from business need to solution deployment
- Understand core data science and analytics philosophy, methodology and tools
- Apply statistical modeling and OR techniques to tackle various aspects of analytics projects
- Make analytics actionable and effectively engage business users and communicate findings

Students will also be exposed to emerging trends, new software tools and best business practices in the DSA area via homework assignments and case studies

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]

References: Bishop, C. *Pattern Recognition and Machine Learning*. Springer, NY: New York, 2007 (ISBN: 0387310738). [Free PDF]

Provost and Fawcett, Data Science for Business - What You Need to Know about Data Mining and Data-Analytic Thinking, O'REILLY, 2013

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, Microsoft Power BI Desktop, MySQL, ILOG, Python, Bash

Grading:

Homework	25%
Mid-term Exam	15%
Final Exam	20%
Semester Project	20%
Participation	10%
Class Exercise/Quiz	10%
[Bonus: Short Bio & Picuture	2% if sent by 9/9]

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

[1] Blackboard website is protected by individual user login names and passwords. The username is the uniquely assigned WSU AccessID. To activate your WSU AccessID or change the password or set an alternate forwarding e-mail address, visit https://computing.wayne.edu/accessid. Call the WSU Computing & Information Technology (C&IT) Help Desk at 313-577-4778 for any difficulties.

(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.

On Working Together: 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 closed book, closed notes 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 two to three students. Students are encouraged (though not required) to pursue a small "research" topic and produce a proceedings paper for a reputable conference or a journal.

Attendance Policy: Students attending any given class are required to join the class within the first five minutes to minimize any class disruptions.

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:

- *The Academic Success Center* (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).
- *The Writing Center* 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. SDS telephone number is 313-577-1851 or 313-577-3365 (TTD only). Once you have your accommodations in place, I will be glad to meet with you privately during my office hours or at another agreed upon time to discuss your needs. 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. Please refer to the SDS website for further information about students with disabilities and the services we provide for faculty and students: 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 Course Schedule:

No. of Classes	Topic
1 Class	Introduction + Course Overview *Course Overview, how data-analytics thinking is transforming different industries and affecting people's life; success stories; graphical showcase, typical machine learning problems (regression, classification, ranking, clustering, dimensionality reduction); typical OR methods (data collection, optimization, simulation)* ***Homework 1: Read assigned papers and book chapters***
1 Class	Install RStudio, get started with R *Walk through the installation, navigation and usage of the GUI and basic commands* *Basics of R data types, data frames and data manipulation methods* *In-class: follow a simple template to work on an example* ***Homework 2: ISLR 2.4 Exercises 1, 6, 8-10***
1 Class	Principles for Data Visualization & Business Analytics Communication *Key points from various books, papers and training materials* *Students to read articles, share finding and experiences and participate in class discussion*
2 Classes	Exploratory Data Analysis and Visualization *Query structured data using SQL; Data wrangling using Excel and R; Visual Analytics with Tableau* *Load data into R from various sources; describe, summarize, subset/slice and plot data* *A pointer to data sources and topics for the team project* ***Homework 3: Visual Analytics with Tableau Desktop***
1 Class	Review of Statistical Concepts, Calculus and Linear Algebra *Probability, random variables, mean, variance, normal distribution, distribution of sample mean, t-distribution* *Continuous Function, Differentiation, Integration, Chain Rule* *Vector, Norm, Inner Product, Matrix, Determinant, Inverse*
3 Classes	Multiple Linear Regression and Model Diagnostics *The importance and desirability of linear models, basic assumptions, coefficient estimation, partition of sums of squares, R-Square, Analysis of Variance* *Multicollinearity, VIF, effect of missing/omitted variables, effect of spread in independent variables* *Variable selection, variable transformation, categorical variable encoding* *Nonlinearity in error terms, Heteroscedasticity, Autocorrelation, and compensation measures* *Regularization* ***Homework 4: ISLR 3.7 & 6.8 Exercises and more***
Mid-term Exam	Closed book, closed note
2 Classes	Working with more forms of data: temporal, geospatial, text and image data *Workshop: Using R and Power BI to explore & visualize data* ***Homework 5: Visualizing GIS data using Microsoft Power BI Desktop***
2 Classes	Logistic Regression *Model formulation, classification and ranking examples in credit scoring and targeted marketing* *Performance metrics: ROC, Lift Chart, c-statistic, precision-discriminability tradeoff for large models* *Interpretation of coefficients, over/under-sampling for imbalanced datasets, bias correction, variable selection* *Case study (using Python and scikit-learn)*
1 Class	Bayesian Decision Process *Maximum likelihood, Naïve Bayesian classifier* ***Homework 6: ISLR 4.7 Exercises and more***
2 Classes	Other Classification Methods *Linear Discriminant Analysis, Decision Trees, Ensemble Methods*
1 Class	Classification workshop (case study of real data) ***Homework 7: Reproducible research with RStudio (R Markdown)***
2 Classes	From Predictive to Prescriptive – Optimization Modeling *Elements and Types of optimization modeling, overview of solution techniques* *Case study (using Python and ILOG)*
2 Classes	Simulation Modeling *Discrete Event Simulation concept, tools and result analysis*

	Homework 8: Solve optimization and simulation problems
1 Class	Analytics Product Development *Production job scripting and scheduling using Bash shell commands* *Components of a successful data analytics product, the Agile philosophy and approach*
4 Classes	Review, team project discussion (scattered across semester), project presentation
Final Exam	Closed book, closed note