WICHITA STATE UNIVERSITY

IME 780AN: Big Data Analytics in Engineering
Spring 2024

Instructor:	Wujun Si, Ph.D.
Department:	Industrial, Systems, & Manufacturing Engineering
Office Location:	120R Engineering Building
Telephone:	(316) 978-5962
Email:	wujun.si@wichita.edu
Method of Contact:	(1) E-mail; (2) Phone (Email always preferred)
Days/Time:	MW: 4:10 PM - 5:25 PM
Office Hours:	MW: 3:00 PM - 4:00 PM
Classroom:	Woolsey Hall 322 (**bring your laptop**)
Prerequisites:	Engineering statistics + good programming skills (at least one programming language, R/Python/Java)
TA:	Saurabh Singh
TA Contact:	sxsingh28@shockers.wichita.edu

Important Dates for this course

Mid-term exam:	**Monday, March 4 (4:10 pm - 5:25 pm)**
Final exam:	**Wednesday, May 1 (4:10 pm - 5:25 pm)**
Final project presentation:	**4/24; 4/29**
Final project report:	**4/29**

For Spring semester 2024, classes begin on Tuesday, January 16, and end Thursday, May 2. The last date to drop a class with 100% refund is Monday, January 29.
No classes: March 11 - 17 (Spring break); April 10.

How to use this syllabus

This syllabus provides you with information specific to this course, and it also provides information about important university policies. This document should be viewed as a course overview; it is not a contract and is subject to change as the semester evolves. If changes are made to the syllabus, an updated version will be posted on Blackboard, and I will announce changes in class and on the Blackboard Announcements tab.

Course Description

This course is designed for graduate students majoring in engineering (with special focuses on industrial, systems and manufacturing engineering). It provides a graduate-level introduction to methods in data science and big data analytics with engineering applications. Specifically, it examines some widely used statistical methods and machine learning tools for engineering data and big data (data with high volume, velocity and variety) analytics. A variety of up-to-date industrial and manufacturing engineering topics are covered as application examples. **R (80%), Python & Java (20%, in last 3 chapters) languages** are used as the major tools for solving your homework and project problems.

Required Textbook

Data Science and Big Data Analytics (Discovering, Analyzing, Visualizing and Presenting Data), EMC Education Services (Editor), Wiley, 2015, ISBN-13: 978-1118876138; ISBN-10: 111887613X.

Reference Tutorials for Coding

Another Book on Data Science: Learn R and Python in Parallel; Nailong Zhang (https://www.anotherbookondatascience.com/)
R for Beginners; Emmanuel Paradis (https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf)

Class Protocol

- In exams all cellphones must be turned off. You cannot use your cellphone as a calculator.
- You are strongly recommended to **have your own textbook** during every class.
- All material and notices will be posted on the Blackboard.
- No web surfing, e-mailing, texting, or any other activity that divides your attention is allowed.

Contact Policy

Although you may attempt to reach me by phone, email communication is always preferred. Feel free to email me any questions or concerns following these guidelines:

- Always use the course name in the subject line of the email
- Remember to sign your name.
- **Always** email me from your WSU email address. Email sent from personal email servers like Gmail, Yahoo, etc., have a tendency to end up in my spam folder, and I never see them. You may also email me through Blackboard via the Email My Instructor tab. I also offer an Ask My Instructor forum on Blackboard which allows common questions to be seen and responded to publicly.
- You should NOT contact me for tech support.
    - Any technical problems involving your computer, or issues regarding file uploading or sharing, should go through the OneStop. You can contact them at 316-978-3909. You can also fill out a request for help form at their website.
    - However, if you have a problem with access or uploading assignments, you *should* let me know before your assignment is due. You will also have to accompany this notification with the file in question, so I can verify that it is completed by the due date/time.

Grading Scale

WSU uses a +/- grading scale for final grades and to calculate grade point averages. In this class, grades are assigned according to the following chart. (Other classes might assign grades differently: Be sure to understand the different grading scales in all of your classes.)

Evaluation	Weight
Homework	20 %
Mid-term Exam	25 %
Final Exam	35%
Project	20%

90 % ≤ A	55 % ≤ C < 67%
87% ≤ A- < 90%	
83 % ≤ B+ < 87%	
80 % ≤ B < 83%	
75 % ≤ B- < 80%	
67 % ≤ C+ < 75%	F < 55%

Homework Assignments

- Homework problems will be assigned and posted on the Blackboard. The assignment due date will be announced in class as well as posted on blackboard (Typically, 1-2 weeks after the posting date).
- Late assignments will not be accepted (if you have special cases, please let me know at least 2 days in advance).

Exams and Course Project

- The midterm (Exam 1) will cover Chapters 1-6. It is open-book, open-notes, and open slides.
- The final exam (Exam 2) will cover Chapters 7 - End. It is open-book, open-notes, and open slides.
- Course project should involve (or at least partially involve) knowledge from Chapter 7-End. Otherwise, a 20% penalty will be applied.

Missed Exams

No makeup exams will be given.

- We only have two exams in this class, i.e., the mid-term and final exams.
- Make sure to plan for the exams accordingly - no makeup exam will be given.

Measurable Student Learning Outcomes

Upon successful completion of this course students will be able to…

- Know basic concepts, definitions and key properties of big data and analytics
  - Identify data structures and data repositories
  - Know state of the practice in analytics
  - Identify key roles for the new big data ecosystem
  - Identify data analytics life cycle
- Understand and implement basic data analytic methods using R
  - Be familiar with the use of R
  - Conduct exploratory data analysis with R
  - Implement statistical methods for evaluation with R
- Understand and implement advanced analytical theory and methods for big data analytics with R & Python:
  - Perform clustering
  - Implement association rules
  - Conduct regression
  - Perform classification
  - Conduct time series analysis
  - Perform text analysis
- Apply technology and tools for advanced analytics: MapReduce and Hadoop
  - Describe analytics of unstructured data
  - Apply the MapReduce and Hadoop paradigm in handling unstructured data
  - Understand Hadoop ecosystem and its application for big data analytics
- Solve real world engineering big data problems with R, Python & Java
  - Identify real world big data problems or projects in engineering
  - Collect (or simulate) "big data" for analytics from multiple industrial/online sources
  - Choose and apply appropriate big data analytics methods to study the problems
  - Solve the problems by implementing the analytics methods in R, Python & Java
  - Utilize the results of analytics in improving engineering design and optimization

Syllabus Policies and Student Resources

All students should familiarize themselves with the course-related policies and student resources that can be found at: **www.wichita.edu/syllabuspolicies**

These include, but may not be limited to:

- COVID-19 Conditions
- Important Academic Dates
- Academic Integrity
- Definition of a credit hour
- Video and Audio recording
- Shocker Alert System

- Intellectual Property
- CARE Team
- Counseling and Prevention Services
- Student Health Services
- Heskett Center and Campus Recreation
- Inclusive Excellence and Respect for Diversity
- First Generation Students
- Names and Pronouns
- Disability Services
- Title IX
- Concealed Carry Policy

Academic Honesty

Students at Wichita State University are expected to uphold high academic standards. WSU will not tolerate a lack of academic integrity. Students are responsible for knowing and following the Student Code of Conduct http://webs.wichita.edu/inaudit/ch8_05.htm and the Student Academic Honesty policy http://webs.wichita.edu/inaudit/ch2_17.htm. When the faculty member determines sanctions are warranted for violations of academic integrity, regardless of severity, the faculty member must report the infraction to the Office of Student Conduct and Community Standards. If you need more information about the process or wish to appeal a decision, please visit https://www.wichita.edu/about/student_conduct/ai.php
Violations of classroom standard include:

- Cheating in any form, whether in formal examinations or elsewhere
- Plagiarism, using the work of others as your own without assigning proper credit to the source
- Misrepresentation of any work done in the classroom or in preparation for class
- Falsification, forgery, or alteration of any documents pertaining to academic records
- Disruptive behavior in a course of study or abusiveness toward faculty or fellow students."

Students with Disabilities

If you are a student with a disability, or believe you might have a disability, which requires accommodations, please contact the Office of Disability Services (ODS) www.wichita.edu/ods to discuss reasonable and appropriate accommodations and eligibility requirements. It is the University's goal that learning experiences be as accessible as possible. If you anticipate or experience physical or academic barriers based on disability ODS will review your concerns and determine, with you, what academic accommodations are necessary and appropriate for you. For example, adaptions of teaching methods, class materials or testing may be made on a case by case basis if warranted, as required by the Americans with Disabilities Act (ADA). All information and documentation of your disability is confidential and will not be released by ODS without your written permission.

Counseling & Testing

The WSU Counseling & Testing Center provides professional counseling services to students, faculty and staff; administers tests and offers test preparation workshops; and presents programs on topics promoting personal and professional growth. Services are low cost and confidential. They are located in room 320 of Grace Wilkie Hall, and their phone number is (316) 978-3440. The Counseling & Testing Center is open on all days that the University is officially open. If you have a mental health emergency during the times that the Counseling & Testing Center is not open, please call COMCARE Crisis Services at (316) 660-7500.

Respect for Diversity

Wichita State University is committed to being an inclusive campus that reflects the evolving diversity of society. To further that goal, Wichita State University does not discriminate in its employment practices, educational programs or activities on the basis of age (40 years or older), ancestry, color, disability, gender, gender expression, gender identity, genetic information, marital status, national origin, political affiliation,

pregnancy, race, religion, sex, sexual orientation, or status as a veteran. Retaliation against an individual filing or cooperating in a complaint process is also prohibited.

Tentative Schedule (may be subject to slight adjustment):

Topic	Date
CHAPTER 1: Introduction to Big Data Analytics	
• 1.1 – 1.4: Big data overview; state of the practice in analytics; key roles for the big data ecosystem; and examples of big data analytics	1/17 1/22
CHAPTER 2: Data Analytics Lifecycle	
• 2.1- 2.8: Data analytics lifecycle overview and Phases 1-6, a case study	1/24
CHAPTER 3: Review of Basic Data Analytic Methods Using R	
• 3.1 – 3.2: Introduction to R	1/29
• 3.2: Exploratory data analysis	1/31
• 3.3: Statistical methods for evaluation	2/5, 2/7
CHAPTER 5: Advanced Analytical Theory and Methods: Association Rules	
• 5.1 – 5.4: Overview; Apriori algorithm; evaluation of candidate rules	2/12
• 5.5 – 5.7: Applications and examples; Validation and testing; diagnostics	2/14
CHAPTER 6: Advanced Analytical Theory and Methods: Regression	
• 6.1: Linear regression	2/19
• 6.2: Logistic regression	2/21, 2/26
• No class, study for Midterm exam	2/28
Midterm Exam (Ch. 1 – 6)	**3/4 (4:10 – 5:25 pm)**
CHAPTER 7: Advanced Analytical Theory and Methods: Classification	
• 7.1: Decision tree	3/6;
• 7.2: Naïve Bayes	3/18; 3/20
CHAPTER 8: Advanced Analytical Theory and Methods: Time Series Analysis	
• 8.1-8.2: Box-Jenkins Methodology; ARIMA Model	3/25 3/27
CHAPTER 9: Advanced Analytical Theory and Methods: Text Analysis	
• 9.1 – 9.4: Text analysis steps, example, raw text collection and representation	4/1
• 9.5 – 9.7: Term frequency, determining sentiments	4/3
CHAPTER 10: Advanced Analytics-Technology and Tools: MapReduce and Hadoop	
• 10.1-10.2: Analytics for Unstructured Data; The Hadoop Ecosystem • Programming beyond R	4/8; 4/15
CHAPTER 11: Selected Topics on Big Data Applications	
• Selected big data applications, programming, and solution • Programming beyond R	4/17 4/22
Project Presentation	4/24; 4/29
Final Exam (Ch. 7 – End)	**5/1 (4:10 – 5:25 pm)**