

IE Technical Elective – Data Analytics in IE

Spring 2021

Course Syllabus

(Subject to change)

Instructor

Changxi Wang, Ph.D. (email: changxi.wang@scupi.cn)

Office: Zone 4-220

Office Hours: Wednesday 13:20-16:25

Teaching Assistant

Jingrao (Gloria) Bian (email: 2017141523024@scu.stu.edu.cn)

Office: TBA

Office Hours: By appointment

Lecture

Wednesday 8:15-11:00; Room: Zone 3-309

Course Description

Data preprocessing, data visualization, classification, clustering, frequent patterns mining, association rules. Project presentation. 3 credit hours.

Course Prerequisites

IE 1070, MATH 280

Course Objectives

1. Learn basics of machine learning models used in industrial engineering
2. Learn to identify industrial problems and formulate them into machine learning problems
3. Learn to use programming languages to analyze data.
4. Learn to recognize and implement various ways of selecting suitable model parameters for different machine learning techniques.

Applicable ABET Outcomes

Students will build

1. An understanding of the data analytics lifecycle.
2. Skills in transformation and merging of data for use in analytic tools.
3. An overview of simple statistical models and the basics of machine learning techniques such as clustering, associations, classification.
4. An understanding of good practices of data science, and conversely recognizing bad practices and why.
5. Skills in the use of tools such as Tableau to explore and mine simple data sets.

Textbook

Jiawei Han, Micheline Kamber, Jian Pei, Data Mining: Concepts and Techniques, Third Edition (The Morgan Kaufmann Series in Data Management Systems) 3rd Edition, 2012

References

数据挖掘导论 Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. Introduction to data mining. Pearson Education India, 2016.

Grading

Homework & exercises, projects, and exam questions related specifically to the objectives above.

Homework & Exercises:	20%
Project:	20%
Mid-Semester Examination:	25%
Final Examination:	<u>35%</u>
	100%

Score	Letter Grade
90.00-100.00	A
85.00-89.99	A-
80.00-84.99	B+
76.00-79.99	B
73.00-75.99	B-
70.00-72.99	C+
66.00-69.99	C
63.00-65.99	C-
61.00-62.99	D+
60.00-60.99	D
0.00-59.99	F

Attendance

There are 16 165-minute lecture periods in the semester. Attendance will be taken for each lecture period. Each student is allowed **two** absences. Each absence, after the second absence, will result in a **1% deduction** from the attendance grade. After the **tenth** absence, the student will not be allowed to take the final exam.

Homework & Exercises

Homework problems and exercises will be assigned periodically and are due as stated. You are each individually required to complete your own exercise sheet which will be collected. Late submission **will not** be accepted. Submissions must be done on **A4 papers and stapled** together at the top left-hand corner. Students' names and ID numbers must be listed on the first page at the top right-hand corner.

Exams

There will two exams, all are CLOSED BOOK, CLOSED NOTES, CLOSED COMPUTER. Students are allowed to bring one A4 page note and it must be hand-written on one side of the paper only. It cannot be a photocopy. If you must miss an exam, you should make alternative arrangements with the instructor before the exam is given. If you miss an exam without prior notification, you will receive a score of "zero" for that exam except under extenuating circumstances.

Make-Up Exams

Students who have not taken both mid-semester and final exams are not eligible for make-up exams. Make-up exams can only be taken by students who have attained between 50.00 % and 59.99 % (out of

100 %) of the total score. Only 75 % of the make-up exam grade can be used to replace the final exam grade. Students taking make-up exams can only attain at most a “D” grade.

Student Opinion of Teaching Surveys

Students in this class will be asked to complete a Student Opinion of Teaching Survey. Surveys will be sent via SCUPI email and appear on your Blackboard landing page during the last three weeks of class meeting days. Your responses are anonymous. Please take time to thoughtfully respond, your feedback is important to me. Read more about Student Opinion of Teaching Surveys.

Avoiding Plagiarism

1. Unacknowledged direct copying from the work of another person, or the close paraphrasing of somebody else's work, is called plagiarism and is a serious offence, equated with cheating in examinations. This applies to copying both from other students' work and from published sources such as books, reports or journal articles.
2. Paraphrasing, when the original statement is still identifiable and has no acknowledgement, is plagiarism. A close paraphrase of another person's work must have an acknowledgement to the source. It is not acceptable for you to put together unacknowledged passages from the same or from different sources linking these together with a few words or sentences of your own and changing a few words from the original text: this is regarded as over-dependence on other sources, which is a form of plagiarism.

Tentative Course Schedule

Lecture	Week	Dates	Topics	Chapter
1	2	March 10	Course Introduction and Review of Syllabus, Introduction to Data Analytics and Data Mining	1
2	3	March 17	Know your data	2
3	4	March 24	Visualization and Distance Measures	2
4	5	March 31	Data Preprocessing Introduction	3
5	6	April 7	Data Preprocessing: PCA & Feature Selection Review Mining Frequent patterns and Associations with Apriori algorithm	6
6	7	April 14	More Frequent patterns and Associations – Apriori continued; Frequent Pattern Growth Method	
7	8	April 21	Classification and introduction to decision trees	8
8	9	April 28	Midterm Exam Review & Midterm Exam	
9	10	May 5	Project Proposal Due today & Tableau visualization software	
10	11	May 12	Classification	8, 9
11	12	May 19	Clustering	10
12	13	May 26	Clustering	11
13	14	June 2	Regression	
14	15	June 9	Project Presentation	
15	16	June 16	Project Presentation	
16	17	June 23	Final Exam Review	
17	18	June 30	Final Exam	