

Instructor: Dr. Ching-yu (Austin) Huang

Email: chuang@kean.edu

Office: NAAB, Room 235

Office Hours: Monday, Wednesday: 12:20–2:50 pm; Tue, Thurs: 12:20–1:50 pm

Website: <http://www.kean.edu/~chuang>

Phone: (908) 737-6157

Class Information:

Course Description: This course will expose students to data mining concepts, techniques, and software utilized in the overall process of discovering knowledge within data.

- **Prerequisites:** Admission to Graduate program.
- **Class:** Monday, Wednesday (11:00 - 12:15pm), NAAB, Room 208
- **Textbook:** Data Mining: Concepts and Techniques, Third Edition by Jiawei Han, Micheline Kamber and Jian Pei, 2012. ISBN: 978-0-12-381479-1.
- **Objectives:** Upon completion of this course, students will be able to:
 - A. Demonstrate an understanding of what knowledge discovery and data mining (KDDM) is.
 - B. Understand the historical perspective of KDDM.
 - C. Expound the differences between KDDM and information access/retrieval.
 - D. Understand KDDM processes, concepts, techniques, algorithms and languages.
 - E. Understand KDDM's applications in business and the sciences.
- **Course Grading:** 2 Midterm (30%), 1 Final Exam (20%), 3 Homework (30%), 1 Research (20%)
A: 100-94, A-: 93-89, B+: 88-84, B: 83-80, B-: 79-76, C+: 75-72, C: 71-68, D: 67-64,
below 64 is failing (a B- or better is needed for CIS graduate program.)
- **Assignments:** Projects and homework assignments will use several computer languages on the dedicated server (Linux).
- **Mid-Term Progress** information will be provided to students by the instructor via KeanWise. Please see the 2018-2019 Undergraduate or Graduate Catalog for more information.

Class Policies:

- No late work will be accepted without permission in advance from the instructor. Late work is penalized.
- No make-up exams are provided without documentation of medical or family emergency. Please arrange your schedule for the next 16 weeks to arrive in class on time and prepared.
- Attendance, participation and lab exercises are very important part of the final grade. Excused absences must be arranged in advance and evidence of the reason for the absence must be provided.
- All work turned in under your name must be your own. Programming assignments must be your own work. No credit will be given for an assignment or homework that is copied – in part or in total – from another person.

Important University Policies and Information:

- The Kean Academic Calendar can be found online at <http://www.kean.edu/KU/Academic-Calendar>
- Tutoring and learning Support services: <http://www.kean.edu/content/tutoring-services>
- The Kean Academic Integrity Policy is online at <http://www.kean.edu/admin/uploads/pdf/AcademicIntegrityPolicy.pdf>. Students are expected to read and abide by the academic integrity guidelines.
- The Kean student code of conduct can be found online at <http://www.kean.edu/KU/Code-of-Conduct>. Students are expected to have read and abide by the code of conduct.
- The Kean Students Rights and Responsibilities handbook is available at <http://www.kean.edu/KU/Forms-Policies-and-Publications>. Students are expected to have read and to follow all regulations.
- All students should register for the university's emergency notification system at <http://www.mir3.com/kean>.
- All students must have a valid Kean email account. If you do not have one, you can request one at <http://www.kean.edu/offices/ocis/forms>; click on E-mail Account Request Form.
- **Title IX** of the Education Amendments of 1972 (Title IX) prohibit discrimination on the basis of sex in education programs or activities. Sexual harassment in any form will not be tolerated at Kean University. Sexual harassment by students should be reported to the Office of Affirmative Action Programs, Office of the Vice President for Student Affairs or the Office of Community Standards and Student Conduct immediately. Information about the University's Sexual Misconduct Policy may be found at the following: <http://www.kean.edu/policies/sexual-misconduct-policy>.
- **Americans with Disabilities Statement & Non-Discrimination Statement:**
Kean University is an affirmative action, equal opportunity institution. Students with documented disabilities who may need special instructional accommodations or who may need special arrangements in the event of an evacuation should notify the instructor as soon as possible, no later than the second week of the term. Students may contact the Office of Disability Services in Downs Hall Room 122 or call 908-737-4910 to discuss special needs.
- **Kean University Non-Discrimination Policy Statement:**
Kean University is an affirmative action, equal opportunity institution.

Week	Date	Topic	Reading	Assignment	
1	1/21/2019	Mon	Martin Luther King Holiday	School closed	
	1/23/2019	Wed	Overview	HW #1 start	1
2	1/28/2019	Mon	Introduction, projects	Research start	2
	1/30/2019	Wed	Introduction, visualization in PHP		3
3	2/4/2019	Mon	Getting to Know Your Data, SQL		4
	2/6/2019	Wed	SQL, aggregation and report		5
4	2/11/2019	Mon	Data Processing - working on text and image		6
	2/13/2019	Wed	Continue, review for midterm #1	HW #1 due by 2/16	7
5	2/18/2019	Mon	President Day Holiday	School closed	8
	2/20/2019	Wed	Midterm #1	HW #2 start	9
6	2/25/2019	Mon	Database vs Data Warehouse		
	2/27/2019	Wed	Data Warehouse and OLAP		10
7	3/4/2019	Mon	Associations and Correlations		11
	3/6/2019	Wed	Mining Frequent Pattern		12
8	3/11/2019	Mon	3/11-17 Spring break	School closed	
	3/13/2019	Wed	3/11-17 Spring break	School closed	
9	3/18/2019	Mon	Mining Frequent Pattern		13
	3/20/2019	Wed	Working on internet data, text	HW #2 due by 3/23	14
10	3/25/2019	Mon	Continue, review for midterm 2	HW #3 start	15
	3/27/2019	Wed	Classification: Basic Concepts		16
11	4/1/2019	Mon	Classification: Basic Concepts		17
	4/3/2019	Wed	Midterm #2		18
12	4/8/2019	Mon	Classification: Advanced methods		19
	4/10/2019	Wed	Classification: Advanced methods		20
	4/12/2019	Fri	Last day to withdraw with 'W' grade	0% refund	
13	4/15/2019	Mon	Cluster Analysis		21
	4/17/2019	Wed	Cluster Analysis	HW #3 due by 4/20	22
14	4/22/2019	Mon	Outlier Detection		23
	4/24/2019	Wed	Outlier Detection		24
15	4/29/2019	Mon	Working on image data		25
	5/1/2019	Wed	Continue, Research presentation	5/4 last day to submit all missing assignments with late penalty	26
16	5/6/2019	Mon	Data Mining trends and impacts		27
	5/8/2019	Wed	Review for Final	Monday schedule	28
17	5/13/2019	Mon	Final exam		29
	5/15/2019	Tue	Review final exam answers		30