

# CSC 7442 – Data Mining and Knowledge Discovery from Datasets

**FALL 2024**

**Louisiana State University  
School of Electrical Engineering and Computer Science  
Division of Computer Science and Engineering**

**Instructor:** Professor Evangelos Triantaphyllou  
**Office:** 3272C Patrick F Taylor (PFT) Hall,  
Phone: (225) 578-1348, E-mail: [etriantaphyllou \(at\) yahoo.com](mailto:etriantaphyllou@yahoo.com)  
**Webpage:** <http://www.csc.lsu.edu/trianta>  
**Classroom:** 1218 Patrick F Taylor (PFT) Hall  
**Meeting Times:** Tuesday & Thursday: from 10:30 to 11:50 AM  
**Office Hours:** **By appointment only (actual office space is under repair).**  
**Please use the Yahoo.com e-mail address given above to contact me.**

**Textbook:** **Introduction to Data Mining**, by Tan, Steinbach and Kumar, published by Addison Wesley, 1<sup>st</sup> edition. Date: May 2005. ISBN-13: 978-0321321367. ISBN-10: 0321321367. and handouts (PDF files and/or URLs).

**Prerequisites:**

Knowledge of elements of graphs, and Boolean algebra. Knowledge of a programming language to write simple programs as part of small projects (any general programming language is fine).

**GOAL:**

The main goal of this course is to introduce the students to the fundamental issues of the fast emerging field of data mining and knowledge discovery from databases. Emphasis is given on the principles that govern this new area of computer science.

**MAIN TOPICS:**

- 1. Introduction**
- 2. The Data**
- 3. Cluster Analysis**
- 4. Other Classification Techniques**
- 5. Boolean function-based Classification Techniques**
- 6. Guided Learning**
- 7. Association Analysis**
- 8. Some Special Topics on Data Mining and Knowledge Discovery from Databases**

<b><u>GRADING:</u></b>	<b>Regular homework assignments:</b>	<b>30%</b>
	(HWs will be given at the end of a teaching unit, i.e., approximately every 1-3 weeks)	
	<b>Small projects:</b>	<b>20%</b>
	(Projects may involve writing simple programs, in any general programming language, to further enhance the learning experience of the subject matter.)	
	<b>One midterm exam:</b>	<b>20%</b>
	<b>Final Exam:</b>	<b>30%</b>
<b>Total points:</b>		<b><u>100%</u></b>
<b>NOTE: All the assignments are INDIVIDUAL.</b>		

**COMMENTS:**

**[1]: The Grading Scale:**

SCORE	GRADE
100, 99, 98	A+
97, 96, 95, 94, 93	A
92, 91, 90	A-
89, 88, 87	B+
86, 85, 84, 83	B
82, 81, 80	B-
79, 78, 77	C+
76, 75, 74, 73	C
72, 71, 70	C-
69, 68, 67	D+
66, 65, 64, 63	D
62, 61, 60	D-
59, or less	F

An F grade (Fail) will be earned for 59 or fewer points. A D grade is not a passing grade for graduate students.

**[2]:** The instructor’s presentations will be highly interactive with lots of Q&A. Asking or answering questions has no impact on your grade. It is absolutely on a voluntary basis. However, make any effort to attend the lectures as there can be more learning opportunities this way. **Always feel free to ask questions during lectures.**

**[3]: No grade curving will take place under any condition.**

**[4]: Some key dates are as follows (based on LSU’s Academic Calendar for 2024-2025):**

August 26, 2024,	Monday:	Classes begin
September 4, 2024,	Wednesday:	Final date for dropping courses without receiving a “W”
September 5, 2024,	Thursday:	Final date for adding courses
October 15, 2024,	Tuesday:	Mid-semester grades are due (by 9:00 am)
October 17-18, 2024,	(Thur. – Frid.):	Fall Holiday
October 20, 2024,	Sunday:	Course scheduling for spring semester begins
November 8, 2024,	Friday:	Final day for dropping courses, 4:30 p.m., deadline
November 27, 2024,	Wednesday:	Thanksgiving Holiday begins, 12:30 p.m.
December 4 – 8, 2024,	(Wed. – Sun.):	Concentrated Study Period
December 7, 2024,	Saturday:	Classes end, 10:00 p.m.
December 13, 2024,	Friday:	Our Final Exam (12:30 – 2:30 p.m.)
December 17-18, 2024,	(Tue. & Wed.):	Final grades are due (9:00 a.m. deadline)
December 19, 2024,	Friday:	Commencement Day
<b>January 13, 2025,</b>	<b>Monday:</b>	<b>Spring Semester 2025 begins!</b>

## Policy Statement Regarding Cheating / Academic Misconduct

A student caught cheating, plagiarizing or violating any of the guidelines described in this Policy Statement will be reported as a case of academic misconduct to the Dean of the Office of Student Advocacy and Accountability (SAA) at LSU.

LSU's Code of Student Conduct, Section 10.1 Academic Misconduct, governs all the work in this course (please check all the materials at: <https://www.lsu.edu/saa/students/codeofconduct.php> ). Unless indicated otherwise, all written work that is handed in must be done only by the individual whose name appears on the document. Regarding student cooperation on the homework assignments, you are encouraged to discuss with other students the homework problems in general terms; however, what you submit must be your own answers. Your instructor and teaching assistant(s) (if TAs have been assigned to this course) are authorized and best qualified to give you help on all work (help will not be given if it provides an unfair advantage). Students are strongly encouraged to carefully study the contents of LSU's Code of Student Conduct on plagiarism and other related issues of academic misconduct. **Thus, you are strongly advised to visit the website of LSU's Office of Student Advocacy and Accountability (SAA) at: <https://www.lsu.edu/saa/>**

For specific information on LSU's definition of **what constitutes academic misconduct**, please visit the following website: <https://www.lsu.edu/deanofstudents/files/lsu-code-of-conduct.pdf>

For some other related concepts of what constitutes plagiarism, visit the following website:  
<http://homeworktips.about.com/od/citationsandbibliography/a/What-Is-Plagiarism.htm>

While you are doing your assignments, **if you use material created by others and found on the web, you must indicate so very clearly.** You must list the source of that material and clearly show which part of your assignment is based on such material from the Web. **Failure to do so will be considered academic misconduct and will be reported.**

**Some instructions while preparing HWs and Projects**

- 1) *Make sure that you understand a given problem before you attempt to answer it. Carefully review the theory covered. Do so by checking the lecture recording (if available), your notes, and ALWAYS read the textbook.*
- 2) *You must solve **all** the problems in a given HW or project by **YOURSELF**.*
- 3) *Pay extra attention in the **organization** of your answers. Solutions should not be mixed up with scratch material in deriving them.*
- 4) *Present and staple your solutions **sequentially**. Scan them by using a scanner or create a PDF file otherwise if possible (no photos please!) and submit them via Moodle.*
- 5) *Make sure that your English is correct, clear and easy to read. Although this is not a composition course, it is very important that you present your solutions in a highly professional and scientific manner.*
- 6) *If you get confused, do not worry. Just relax, organize your thoughts, and try to see the problem from a simple, but still accurate, point of view.*
- 7) *You should have no doubt that **you** (assuming you have studied adequately) are capable of answering all the problems in the HWs and projects!*

**A hint...**

Next is the first problem which would have been given to you in the midterm and final exams (but we will do not have such exams this time) along with the correct answer. It is usually assigned 5 points.

\*\*\*\*\*

**PROBLEM 1: (5 points)**

What is the single most important step in solving any science / engineering problem?

**Answer:** To define the problem **correctly** and **enjoy** the solution process.

(You must underline the right words to earn all 5 points.)