WEBSTER UNIVERSITY • WEBSTER GROVES, MO • GEORGE HERBERT WALKER SCHOOL OF BUSINESS & TECHNOLOGY • MATHEMATICS/COMPUTER SCIENCE

DATA MINING METHODS

MATH-3220 Section 01
3 Credits FA 2014
25/08/2014 to 19/12/2014 Modified 27/07/2014

MEETING TIMES

Seminar

Monday, Wednesday, Friday, 13:00 to 13:50, EAB

CONTACT INFORMATION

John Aleshunas

Email: jalesh@webster.edu
Office: EAB 376
Phone: 314-246-7565
Website: http://mercury.webster.edu/aleshunas

DESCRIPTION

This course surveys the current techniques of problem solving using modern heuristics. It covers classic methods of optimization, including dynamic programming, the simplex method, and gradient techniques, as well as recent innovations such as simulated annealing, tabu search, and evolutionary computation. Besides exploring a compendium of specific techniques, this course also delves into the approaches of framing and attacking the issue of problem solving itself. Students will present their findings and recommendations in written and oral project reports.

Requisites

MATH 1610 Calculus I

OBJECTIVES

Course Objectives

1. To develop in students good problem organization and solving skills.
2. To teach students, through experience accompanied by constructive feedback, to

communicate their ideas orally in an organized and effective manner.
3. To teach students, through experience accompanied by constructive feedback, to communicate their ideas in writing in an organized and effective manner.
4. To develop in students the skill of critical analysis and critical thinking.
5. To familiarize students with the mathematical techniques and methods used in data mining.
6. To teach students how to apply data mining algorithms to solve practical problems.

OUTCOMES

At the completion of this course, each student will be able to:

1. Demonstrate good problem organization and solving skills.
2. Communicate their ideas orally in an organized and effective manner in classroom discussion.
3. Communicate their ideas in writing in an organized and effective manner using Project Reports.
4. Choose and apply mathematical algorithms and heuristics that are appropriate to given problems or desired outcomes.
5. Describe the mathematical algorithms and heuristics used in solving a specific problem and explain what problems they are best suited for.
6. Conduct a research project using a chosen mathematical algorithm or heuristic technique and present a formal report on the outcome.

REQUIRED TEXTBOOK

No Textbook Required

MATERIALS

Additional Supporting Content

Instead of a required textbook, the instructor will provide additional materials to augment the course texts with current published research.

DELIVERABLES

Exercises and Reports

The course exercises will help you develop working knowledge and skills in research activities. Each exercise will incrementally introduce new capabilities and reinforce previously introduced concepts. The overall goal of these exercises is to prepare you for your course project work.

Research Project

The individual research project provides you the opportunity to experiment with a selected data mining topic. You may select any research topic, subject to my approval. Remember, we are trying to gain competency in these techniques and algorithms and some areas, for example, where the problem domain is not constrained and well understood, may not be as productive as others. Additionally, I want to expose you to a variety of data mining topics.

You can will develop a working experiment with a data mining technique or algorithm and publish your findings in a report. You will conclude your project with a presentation in finals week.

This is a formal paper, and it requires a formal presentation. This is an opportunity for you to share your work with the class. Plan for a ten-minute presentation before questions and comments. Don’t read your paper. Determine the most important and interesting parts of your paper for the presentation (three items at most). It is not necessary to include in your presentation everything in the paper, and in fact, there will not be time to do so. You may use overhead transparencies or other aids in you presentation.

EVALUATION

Criteria

The UNDERGRADUATE catalog provides these guidelines and grading options:

- **A, A-** superior work in the opinion of the instructor
- **B+, B, B-** good work in the opinion of the instructor
- **C+, C, C-** satisfactory work in the opinion of the instructor
- **D+, D** passing, but less than satisfactory work in the opinion of the instructor
- **I** incomplete work in the opinion of the instructor
- **ZF** An incomplete which was not completed within one year of the end of the course

- **F** unsatisfactory work in the opinion of the instructor; no credit is granted
- **W** withdrawn from the course
- **IP** course in progress
- **NR** not reported for the course
- **Z** a temporary designation given by the registrar indicating that the final grade has not been submitted by the instructor. When the final grade is filed in the Office of the Registrar, that grade will replace the Z.

Your grade will be compiled from each of the class evaluation components in the following proportions:

Exercises and Reports	60%
Research Project	25%
Attendance and Participation	15%
Total	100%

The course grading scale is:

93 to 100%	A
90 to 92%	A-
87 to 89%	B+
83 to 86%	B
80 to 82%	B-
77 to 79%	C+
73 to 76%	C
70 to 72%	C-
60 to 69%	D
Below 60%	F

COURSE POLICIES

Attendance

Attendance will be taken at each class meeting. Attendance is required. Please notify me in advance of schedule problems. You will be responsible for all material covered in class as well as in the textbook. If you are absent, you should make arrangements with another student for class notes and with me for any class handouts. **Excessive absences will reduce a student's grade for the course**.

Assignments

Assignments are due at the beginning of class. Reading assignments should be completed **before** class on the related topic to enable you to bring questions to class and gain more from the class discussions.

INSTITUTIONAL POLICIES

University policies are provided in the current course catalog and course schedules. They are also available on the university website. This class is governed by the university's published policies. The following policies are of particular interest:

Academic Honesty

The university is committed to high standards of academic honesty. Students will be held responsible for violations of these standards. Please refer to the university's academic honesty policies for a definition of academic dishonesty and potential disciplinary actions associated with it.

Drops and Withdrawals

Please be aware that, should you choose to drop or withdraw from this course, the date on which you notify the university of your decision will determine the amount of tuition refund you receive. Please refer to the Add/Drop/Withdraw section of the academic catalog for further information and to find the deadlines for dropping a course with a full refund and for withdrawing from a course with a partial refund.

Special Services

If you have registered as a student with a documented disability and are entitled to classroom or testing accommodations, please inform the instructor at the beginning

of the course of the accommodations you will require in this class so that these can be provided.

Disturbances

Since every student is entitled to full participation in class without interruption, disruption of class by inconsiderate behavior is not acceptable. Students are expected to treat the instructor and other students with dignity and respect, especially in cases where a diversity of opinion arises. Students who engage in disruptive behavior are subject to disciplinary action, including removal from the course.

Grading

Please refer to the most recent academic catalog for information on the Webster University grading policy.

Student Assignments Retained

From time to time, student assignments or projects will be retained by The Department for the purpose of academic assessment. In every case, should the assignment or project be shared outside the academic Department, the student's name and all identifying information about that student will be redacted from the assignment or project.

Contact Hours for this Course

It is essential that all classes meet for the full instructional time as scheduled. A class cannot be shortened in length. If a class session is cancelled for any reason, the content must be covered at another time.

SCHEDULE

When	Topic	Notes
Week 1	Introduction	Introduction Terminology Problem Solving & Planning
Week 2	Concepts	Algorithms vs. Heuristics

When	Topic	Notes
		Defining the Problem Models and Basic Techniques
Week 3	Traditional Methods	Traditional Methods Exhaustive Search Local Search
Week 4	Traditional Methods	Greedy Algorithms Dynamic Programming Branch and Bound Techniques
Week 5	Improving the Search	Escaping Local Optima Simulated Annealing Tabu Search
Week 6	Evolutionary and Genetic Algorithms	Evolutionary and Genetic Algorithms
Week 7	Evolutionary and Genetic Algorithms	Evolutionary and Genetic Algorithms
Week 8	An Example	The Traveling Salesman Problem [TSP]
Week 9	Constrained Search	Constraint Handling Techniques
Week 10	Improving the Search - Revisited	Tuning Algorithms and Heuristics to the Problem
Week 11	Noisy Data and Data Cleansing	Noisy Data and Data Cleansing
Week 12	Neural Networks	Biological Foundations

When	Topic	Notes
		Activation Functions Learning Optimization
Week 13	Machine Learning	Machine Learning Learning and Memory Models Supervised Learning Algorithms
Week 14	Accomodating Uncertainty	Fuzzy and Hybrid systems
Week 15	Accomodating Uncertainty	Fuzzy and Hybrid systems
Week 16	Exam Week	**EXAM WEEK** Research Reports