

FLORIDA ATLANTIC UNIVERSITY™

Graduate Programs—NEW COURSE PROPOSAL¹

UGPC APPROVAL _____
 UFS APPROVAL _____
 SCNS SUBMITTAL _____
 CONFIRMED _____
 BANNER POSTED _____
 CATALOG _____

DEPARTMENT: ITOM

COLLEGE: BUSINESS

RECOMMENDED COURSE IDENTIFICATION:

PREFIX ISM COURSE NUMBER 6127 LAB CODE (L or C) _____

(TO OBTAIN A COURSE NUMBER, CONTACT MJENNING@FAU.EDU)

COMPLETE COURSE TITLE: DATA MINING & DATA WAREHOUSING

EFFECTIVE DATE

(first term course will be offered)

_____ SPRING 2015 _____

CREDITS²: 3

TEXTBOOK INFORMATION: DATA MINING FOR BUSINESS INTELLIGENCE, SECOND EDITION;
 GALIT SHMUELI, NITIN R. PATEL AND PETER C. BRUCE
 ISBN: 978-0-470-52682-8 JOHN WILEY & SONS; 2010

GRADING (SELECT ONLY ONE GRADING OPTION): REGULAR V SATISFACTORY/UNSATISFACTORY _____

COURSE DESCRIPTION, NO MORE THAN THREE LINES:

INTRODUCES THE CORE CONCEPTS OF DATA MINING (DM), ITS TECHNIQUES, IMPLEMENTATION, AND BENEFITS. COURSE ALSO IDENTIFIES INDUSTRY BRANCHES THAT MOST BENEFIT FROM DM, SUCH AS RETAIL, TARGET MARKETING, FRAUD PROTECTION, HEALTH CARE AND SCIENCE, AND WEB AND E-COMMERCE. DETAILED CASE STUDIES AND USING LEADING MINING TOOLS ON REAL DATA ARE PRESENTED.

PREREQUISITES*: NONE

COREQUISITES*: NONE

REGISTRATION CONTROLS (MAJOR, COLLEGE, LEVEL)*: NONE

* PREREQUISITES, COREQUISITES AND REGISTRATION CONTROLS WILL BE ENFORCED FOR ALL COURSE SECTIONS.

MINIMUM QUALIFICATIONS NEEDED TO TEACH THIS COURSE: TERMINAL DEGREE (PH.D.) WITH QUANTITATIVE ANALYSIS COURSE WORK

Faculty contact, email and complete phone number:
 Mary Schindlbeck, Ph.D.
 mschind2@fau.edu
 561-297-3661

Please consult and list departments that might be affected by the new course and attach comments.³ NO DEPARTMENTS ARE AFFECTED.

Approved by:

Department Chair: _____

College Curriculum Chair: _____

College Dean: _____

UGPC Chair: _____

Graduate College Dean: _____

UFS President: _____

Provost: _____

Date:

9/3/14
9-8-2014
9-8-2014
10/8/14
10-15-14

1. Syllabus must be attached; see guidelines for requirements: www.fau.edu/provost/files/course_syllabus.2011.pdf

2. Review Provost Memorandum: Definition of a Credit Hour www.fau.edu/provost/files/Definition_Credit_Hour_Memo_2012.pdf

3. Consent from affected departments (attach if necessary)

Email this form and syllabus to UGPC@fau.edu one week before the University Graduate Programs Committee meeting so that materials may be viewed on the UGPC website prior to the meeting.



Florida Atlantic University
COLLEGE OF BUSINESS

ISM 6127

CRN #

Data Mining & Data Warehousing

Fall 2014

FL 411

Tuesday 7:10pm – 10:00pm

Professor Information

Mary Schindlbeck, Ph.D.

Boca Raton Campus - FL 317

mschind2@fau.edu

561-297-3661

Office Hours

Tuesday 1:00pm- 2:00pm and 4:00pm-7:00pm FL317

Thursday 12:00pm-2:00pm FL317

Required Text and Materials

Text: Data Mining for Business Intelligence, Second Edition;

a new text includes a 6-month license code for XLMiner the required software.

Galit Shmueli; Nitin R. Patel and Peter C. Bruce

ISBN: 978-0-470-52682-8 John Wiley & Sons; 2010

Data Mining Software: XLMiner® for Windows - a comprehensive data mining add-in for Excel, with neural nets, classification and regression trees, logistic regression, linear regression, Bayes classifier, K-nearest neighbors, discriminant analysis, association rules, clustering, and principal components analysis. Students enrolled in the class will have access to the XLMiner software in the instructional lab FL411. A new text includes an access code for a 6-month license for XLMiner which you can download onto your personal computer.

You can read more about XLMiner on the tool's web site: <http://www.resample.com/xlminer/>

Students should have a working knowledge of basic math (algebra) and Microsoft Excel. Students should have access to Excel spreadsheet software and are assumed to be familiar at an intuitive level with general business practices of collecting, storing and using data.

Course Description

Introduces the core concepts of data mining (DM), its techniques, implementation, and benefits. Course also identifies industry branches that most benefit from DM, such as retail, target marketing, fraud protection, health care and science, and web and e-commerce. Detailed case studies and using leading mining tools on real data are presented.

Course Prerequisites, Credit Hours, and Class Time Commitments

No course prerequisites

3 Credit Hours

According to Florida State Statute 6A-10.033, students must spend a minimum 37.5 of in class time during a 3-credit course. Additionally, students enrolled in a 3-credit course are expected to spend a minimum of 75 hours of out-of-class-time specifically working on course-related activities (i.e., reading assigned pieces, completing homework, preparing for exams and other assessments, reviewing class notes, etc.) and fulfilling any other class activities or duties as required.

Course Learning Objectives

The primary aim of this course is to acquaint students with the practical challenges that are encountered when solving real-world data mining problems. By allowing students to encounter the complete picture of going from business level problem formulation, through the analysis of quantity and quality of available data, as well as selection of suitable algorithms and evaluation of the obtained results. Students are expected to learn not only about the advantages of various data mining methods, but also about their limitations. The main focus in this course will be on applying, in practice, knowledge and concepts from business models and on understanding the applicability of data mining.

Students will reinforce the learning of data mining concepts by means of data analysis techniques to make better business decisions through proper data preparation, data exploration and tools for solving data mining problems. Students will be introduced to advanced concepts such as data mining applications, data warehouses, web mining, text mining, and ethical aspects of data mining. Students will learn to mine heterogeneous data and demonstrate proficiency in classification and prediction applications such as neural networks, linear regression, cluster analysis, market basket analysis and decision trees.

Working individually, students will demonstrate proficiency in applying data mining analytical techniques on an advanced real world business problem that examines a large amount of data to discover new information in addition to analyzing and evaluating technique effectiveness with evolving technologies by presenting a self-designed semester project. Commencing with several singular technique projects and concluding with the comprehensive semester project, students will reinforce their oral skills by way of presentations as well as written and critical thinking skills by the use of executive memos and a final research paper requiring quantitative analysis and evaluation.

Grading Scale

| | | | |
|----|------------|----|-----------|
| A | 93.00-100% | C | 73-76.99% |
| A- | 90-92.99% | C- | 70-72.99% |
| B+ | 87-89.99% | D+ | 67-69.99% |
| B | 83-86.99% | D | 63-66.99% |
| B- | 80-82.99% | D- | 60-62.99% |
| C+ | 77-79.99% | F | < 60 % |



Course Evaluation Method

- Five Team Projects (each 5%) – 25%
- Data Mining Discussions – 10%
- Midterm Exam– 20%
- Final Exam - 20%
- Individual Research Paper & Presentation - 25%

Exams: The two exams will include multiple choice questions and data mining application problems, administered on Blackboard during class and will **cover content from the text, material presented in the lectures and practical lessons from the assignments.** Usually, students will be asked to interpret results from applying a specific data mining method, such as confusion matrices and classification false positive/negative rates. Therefore, assignments, discussions, class attendance and good note taking are essential elements for success. Each exam has a time limit of 90 minutes.

Team Projects: Projects will enforce a specific data mining method or principle. The team should be of exactly 2 students. Finding a team partner is solely students' responsibility. Choose your partner carefully, identify if your goals in this course are common and if the level of commitment is the same. If there are differences on these two basic criteria, chances are you will not collaborate effectively and there will be problems down the road. It is to your best advantage to document (email) your communications to avoid complications; if you feel more comfortable, use Google Docs to complete your projects and/or feel free to cc: your emails to me. If there is an odd number of the students in class, the instructor will have the discretion to place the remaining student in a team whose team members should do everything possible to work together as a team of three.

Your team will use the same data set for each of the first five projects unless specified to do otherwise. For each assignment you will post all of the files you created in the Assignment Section of Black Board before the due date and time; penalty of 10% for each day exists for late submissions. Some teams will present their findings and other teams will participate in a discussion about the findings; our class will be similar to a project team. No individual assignments will be accepted for the first five projects.

While you will collaborate with your peers on the 5 team projects, each team member must contribute an equal portion of effort or work towards each phase of the project. Based on peer evaluations and observations of group work by the professor, team members contributing more than their fair share may receive higher grades than other group members. It is the group's responsibility to advise the professor of any issues or challenges that may hinder your ability to complete any phase of the project.

For each project submitted the file names will contain assignment number, dataset name and the last names of the team members. **For example, for assignment 1 using the insurance dataset, the file name for students Jane Smith and Joe Cole should be ASG1_Insurance_Smith_Cole.xxx** (.docx or .xlsx depending on the type of file). Worksheets in an Excel workbook must be appropriately named and unnecessary sheets must be removed.

Project Submissions:

The project submission must include the following:

- The actual Excel spreadsheet file(s) where the method/tool was applied.
- The necessary additions such as confusion matrices, classification rates, etc., that reinforce the appropriate conclusions (can be added as worksheets to the original Excel file).
- Memorandum that concisely presents, summarizes, and analyzes the results (draw meaningful conclusions and comparisons). Detailed information about the memo and some examples can be found on blackboard.

Individual Research Paper & Presentation:

A detailed document and rubric is provided on Blackboard regarding all requirements of the final data mining project. A research project proposal including the data source and data description must be pre-approved by the instructor by the proposal due date.

The project will require locating a large data set (between 3,000 and 10,000 records) with at least 8 variables of differing data types. The goal of the final project is to go through the full data mining cycle with respect to a particular data set (including the specification of the business problem to be solved, the specification of the data mining tasks to be performed, selection, preprocessing, integration, and transformation of the data, application of several data mining tasks and the discovery of patterns, evaluation of patterns, and recommending specific actions with respect to relevant findings).

This project will demonstrate a comprehensive understanding of the course.

Research Paper: The goal of the "research project" is to go beyond the class material and examine one of the data mining topics in a more in-depth manner. The evaluation of the paper will be based on thoroughness (including adequate coverage of relevant issues/techniques as well as references to related work), soundness (including justification for any claims made, illustrative example, correctness and adequate analysis of connections or relationships among concepts or techniques) and organization. Paper should include the following sections: executive summary (exploratory data analysis and mapping to the problem), data analysis, modeling approach (rationale for model selection and evaluation process), results and conclusions (interpretation and implications).

Presentation: Final project will be presented to a peer audience and not exceed 20 minutes. The presentation will include the analysis of each technique as well as a comparison/contrast of the techniques applied. Discussion of the output can be included in PowerPoint slides and should be able to be demonstrated. Complete final project rubrics are available on Blackboard.

Participation & Discussion:

The team assignments, after submission, will be discussed in a class session. Far from everything will be clear and exact in these sessions – we will need a lot of input and brainstorming – a normal process when engaged in highly analytical work such as data mining and cleaning the data. Students are expected to actively participate and generate discussions on the techniques used and the results. Open discussions and participation are important elements of the course. Whether your techniques, methods and conclusions are correct or wrong, the discussion grade will not be affected. The goal is to



reach the best method and solution through sharing each group's ideas. Participation also includes bringing relevant current topics in the news into the classroom.

The course will require attending one 2 hour and 50 minute lecture with labs each week and approximately 4 to 5 hours of outside work to learn the material through readings, tutorial lessons and selected problems for each class session.

Additional Course Policies

Missed Exams

It is important that each exam be taken at the scheduled time and date. Any excusable absence (official athletic event, religious holiday, etc.) must be documented by a verifiable source and I must be notified at least one week prior to the exam. If you are absent from an exam due to illness or emergency, you must notify me by e-mail within 24 hours of the missed exam and provide verifiable documentation within one week of the exam date; the make-up policy is not applicable if you fail to report an absence as stated above. There will be two semester exams, each covering approximately one-half of the course material. A mid-term exam missed with prior documented approval as stated above may be made up by the Final exam. The score earned on the Final exam will be used for both the final and for the missed exam. An exam missed without prior approval and verifiable documentation that the unapproved absence was unavoidable as stated above cannot be made up.

Late Assignments

Grade penalty equal to 10 percent of the project grade per day late will be applied after the project's due date.

Attendance Policy

Learning is an interactive process and success in this course depends on the experiences the students bring to the classroom (our learning community). Therefore attendance is an important aspect of this course. Attendance will not be taken. However, you are responsible for everything that takes place in class. Additional homework assignments, their due dates, and changes to the tentative schedule will be announced in class. Occasionally, unannounced in-class exercises (or quizzes) will be given; if missed, these cannot be made up. Due to the cumulative nature of the material it is imperative that students keep up with the course materials on a daily basis. Attendance is strongly suggested and is a prerequisite for successful completion of this course. Missing classes will adversely affect your performance. The probability of successfully passing the tests in the course is directly dependent on regular attendance, studying the assigned materials and completing projects and lab exercises in a timely manner.

Etiquette and/or Netiquette Policy

Each student is responsible for keeping up with the class schedule, checking your FAU email account, and



Florida Atlantic University
COLLEGE OF BUSINESS

checking the course Blackboard site on a regular basis. If you use a non FAU email address as your primary address, arrange for FAU email to be forwarded.

Please use the Message option in Blackboard to contact me. Do not use e-mail.

Anti-plagiarism Software

Written components of any assignment or project may be submitted to anti-plagiarism software to evaluate the originality of the work. Any students found to be submitting work that is not their own will be deemed in violation of the University's honor code discussed below.



Florida Atlantic University
COLLEGE OF BUSINESS

Tentative Course Schedule

| Date | Lecture | Readings | Assignments & Due Dates |
|---------------|--|--|----------------------------------|
| 8-19 | Introduction to Data Mining & Data Mining Techniques Data Preprocessing | Shmueli-Chapter 1 Shmueli-Chapter 2 | |
| 8-26 | Data Exploration Data Visualization | Shmueli -Chapter 3 BB-XLMiner Notes | Install XLMiner |
| 9-2 | Data Reduction Evaluation of Model Performance PRESENT & DISCUSS – Data Exploration | Shmueli -Chapter 4 Shmueli -Chapter 5 | Assignment 1 Data Exploration |
| 9-9 | Prediction & Classification Linear Regression & Logistic Regression Regression Lab: XLMiner and Excel | Shmueli -Chapter 6 Shmueli -Chapter 10 BB-Regression Notes | |
| 9-16 | PRESENT & DISCUSS – Regression | | Assignment 2 Regression |
| 9-23 | Classification & Regression Trees Decision Tree Lab: XLMiner | Shmueli -Chapter 9 BB-Decision Tree Notes | |
| 9-30 | PRESENT & DISCUSS - Decision Trees | | Assignment 3 Decision Tree |
| 10-7 | MIDTERM EXAM | | |
| 10-10-14 | Last day to drop or withdraw without receiving an F in the course. | | |
| 10-14 | Cluster Analysis K-Means/Clusters Lab: XLMiner | Guest Speaker Shmueli -Chapter 14 BB-K-Means Notes | FINAL PROJECT PROPOSAL Due |
| 10-21 | PRESENT & DISCUSS - K-Means Clustering Neural Networks in Data Mining Neural Networks Lab: XLMiner | Shmueli -Chapter 11 BB-NN Notes | Assignment 4 K-Means |
| 10-28 | PRESENT & DISCUSS - Neural Networks Association Rules - Market Basket Analysis | Shmueli -Chapter 13 BB-MBA Notes | Assignment 5 Neural Network |
| 11-4 | FINAL PROJECT Presentations | | |
| 11-11 | Tuesday - Veteran's Day | | |
| 11-18 | FINAL PROJECT Presentations | | Final Project Reports |
| 11-25 | FINAL PROJECT Presentations | | |
| 11-27 - 11-30 | Thanksgiving Recess | | |
| 12-1 - 12-3 | Reading Days | | |
| 12-9 | FINAL EXAM 7:00pm | | |

*Shmueli - Course Textbook BB – Blackboard

Online Resources and References Material List

JOURNALS

ACM Transactions on Knowledge Discovery in Data (TKDD).

SIGKDD Explorations, a magazine of the SIGKDD, the data miners professional group

Data Mining and Knowledge Discovery journal published by Springer

Journal of Big Data, a SpringerOpen Journal.

Journal of Data Mining and Knowledge Discovery, Bioinfo publications, India.

Journal of Data Science, international journal devoted to applications of statistical methods at large

International Journal of Data Mining & Knowledge Management Process

Journal of Machine Learning Research

BOOKS

J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, ISBN-13: 978-0123814791

I.H. Witten, E. Frank, M.A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, ISBN-13: 978-0123748560

F. Provost and T. Fawcett, *Data Science for Business: What you need to know about data mining and data-analytic thinking*, ISBN-13: 978-1449361327

Selected University and College Policies

Code of Academic Integrity Policy Statement

Students at Florida Atlantic University are expected to maintain the highest ethical standards. Academic dishonesty is considered a serious breach of these ethical standards, because it interferes with the university mission to provide a high quality education in which no student enjoys an unfair advantage over any other. Academic dishonesty is also destructive of the university community, which is grounded in a system of mutual trust and places high value on personal integrity and individual responsibility. Harsh penalties are associated with academic dishonesty. For more information, see University Regulation 4.001.

Disability Policy Statement

In compliance with the Americans with Disabilities Act (ADA), students who require special accommodation due to a disability to properly execute coursework must register with the Office for Students with Disabilities (OSD) – in Boca Raton, SU 133, (561) 297-3880; in Davie, MOD 1, (954) 236-1222; in Jupiter, SR 117, (561) 799-8585; or, at the Treasure Coast, CO 128, (772) 873-3305 – and follow all OSD procedures.



Florida Atlantic University COLLEGE OF BUSINESS

Religious Accommodation Policy Statement

In accordance with rules of the Florida Board of Education and Florida law, students have the right to reasonable accommodations from the University in order to observe religious practices and beliefs with regard to admissions, registration, class attendance and the scheduling of examinations and work assignments. For further information, please see Academic Policies and Regulations.

University Approved Absence Policy Statement

In accordance with rules of the Florida Atlantic University, students have the right to reasonable accommodations to participate in University approved activities, including athletic or scholastics teams, musical and theatrical performances and debate activities. It is the student's responsibility to notify the course instructor at least one week prior to missing any course assignment.

College of Business Minimum Grade Policy Statement

The minimum grade for College of Business requirements is a "C". This includes all courses that are a part of the pre-business foundation, business core, and major program. In addition, courses that are used to satisfy the university's Writing Across the Curriculum and Gordon Rule math requirements also have a minimum grade requirement of a "C". Course syllabi give individualized information about grading as it pertains to the individual classes.

Incomplete Grade Policy Statement

A student who is passing a course, but has not completed all work due to exceptional circumstances, may, with consent of the instructor, temporarily receive a grade of incomplete ("I"). The assignment of the "I" grade is at the discretion of the instructor, but is allowed only if the student is passing the course.

The specific time required to make up an incomplete grade is at the discretion of the instructor. However, the College of Business policy on the resolution of incomplete grades requires that all work required to satisfy an incomplete ("I") grade must be completed within a period of time not exceeding one calendar year from the assignment of the incomplete grade. After one calendar year, the incomplete grade automatically becomes a failing ("F") grade.

Withdrawals

Any student who decides to drop is responsible for completing the proper paper work required to withdraw from the course.



Grade Appeal Process

A student may request a review of the final course grade when s/he believes that one of the following conditions apply:

- There was a computational or recording error in the grading.
- Non-academic criteria were applied in the grading process.
- There was a gross violation of the instructor's own grading system.

The procedures for a grade appeal may be found in Chapter 4 of the University Regulations.

Disruptive Behavior Policy Statement

Disruptive behavior is defined in the FAU Student Code of Conduct as "... activities which interfere with the educational mission within classroom." Students who behave in the classroom such that the educational experiences of other students and/or the instructor's course objectives are disrupted are subject to disciplinary action. Such behavior impedes students' ability to learn or an instructor's ability to teach. Disruptive behavior may include, but is not limited to: non-approved use of electronic devices (including cellular telephones); cursing or shouting at others in such a way as to be disruptive; or, other violations of an instructor's expectations for classroom conduct.

Faculty Rights and Responsibilities

Florida Atlantic University respects the right of instructors to teach and students to learn. Maintenance of these rights requires classroom conditions which do not impede their exercise. To ensure these rights, faculty members have the prerogative:

- To establish and implement academic standards
- To establish and enforce reasonable behavior standards in each class
- To refer disciplinary action to those students whose behavior may be judged to be disruptive under the Student Code of Conduct.