

MSBA 206 – Data Mining for Business

Instructor: Bang Tran, Assistant Professor in the Department of Computer Science

Office: RVR 5042

Phone: (916)-278-6088

Email: message through Canvas or s.tran@csus.edu
(Please include “MSBA 206” at the beginning of the subject line)

Website <https://webpages.csus.edu/s.tran/>

Office Hours: Thursday, 2 p.m.– 5:00 p.m.

<https://csus.zoom.us/j/84415087613>

Textbook: Data Mining For Business Analytics Concepts, Techniques And Applications in Python by Galit Shmueli, Nitin R. Patel, and Peter C. Bruce, 1st Edition.

As a plus to the textbooks, course slides, assignments, and other artifacts will be available to you through Canvas.

Course Content:

This course covers data mining methods commonly used in business including advanced data visualization, dimension reduction, classification methods, association rules, cluster analysis, and forecasting time series analysis. This course focuses on applications of data mining methods in the business environment.

Prerequisite:

MSBA 205

Course Policies:

Tentative Grading Policy (subject to change during the semester):

Middle term exam + Final exam	30% + 35%
Assignments	35%

Grading Breakdown (%):

A = 93-100	C = 73-76
A- = 90-92	C- = 70-72
B+ = 87-89	D+ = 67-69
B = 83-86	D = 63-66
B- = 80-82	D- = 60-62
C+ = 77-79	F = 59 or below

Students are required to keep backup (machine-readable) copies of all submitted work, and also to keep all returned (graded) work, until after final grades are posted.

All the assignments will be graded with 100 as the highest. The final scores will then be the weighted score and rounded up to match the above scale. Please note that the final score is not negotiable. Also, the highest grade in the university system is an A. You will need a passing grade for all course exams to pass the whole class.

For labs and projects, as long as your answer involves certain commands or operations in specific software, screenshots are also needed to demonstrate your result.

The lab report, project report, and oral presentation are supposed to illustrate or explain **what you did** (commands or configurations) and **what you got** (screenshots and analysis). They will all be evaluated based on the following grading criteria.

Correctness	25%
Completeness	25%
Clearness	25%
Quality of English writing	25%

Individual Work

All the parts of this course should be accomplished **independently!**

Note: Specific instructions for labs and projects, including the deliverable requirements and due dates, will be assigned in separate documents after corresponding lectures. So, please make sure you get this important information in class or via Canvas/email. Programming exercises will be graded for appearance, programming style, and comments as well as for correctness. All output should be identified and illustrated, and the input used for any program should be listed and explained. Your programming assignments will be completed on a Linux system (or Windows if possible). When applicable, input should be read from a file/console, and output redirected to a file/console, so that the inputs, outputs, and program listings can be easily printed/snapshotted. Unlike labs and projects, homework assignments will not be graded, but it is a good idea to do them because they will help you in the quizzes/exams.

Secure Programming Project: This semester socket programming projects will be graded for **correctness** (including functionality completeness and programming clearness in terms of appearance, programming style, and comments), which occupies 60% points, as well as for **robustness**, which earns the other 40% points. All output should be identified and illustrated, and the input used for any program should be listed and explained. When applicable,

input should be read from a file/console, and output redirected to a file/console, so that the inputs, outputs, and program listings can be easily printed/snapshots.

Submission Rules:

Each submission needs to be in an **electronic version** (through Canvas). Any file needs to be named according to one of the following formats (depending on the submission type). **Please do NOT use the .txt format. Word** format is preferable. Please also write your class **section number** in the document (otherwise, you will lose points).

MSBA206_assignment#_name,

For example, if James Green is submitting project 1, the file name of the submission should be MSBA206_assignment1_James_Green. **On the first page of each submitted document, please always list your name as the author.** Please note: if the attachment is not according to the proper format as stated above, it will not be accepted.

Due Date and Late Submission:

Please see the above due dates specified in the instruction documents on Canvas from the instructor. The mid-term exam will be performed respectively at the second class of week 7.

Late submissions will be penalized by the following rules:

- Second-day submission: **20% off** the assignment grade.
- Third-day submission: **50% off** the assignment grade.
- After the third day: **100% off** the assignment grade.

Laptop and Cell Phone Regulation:

A laptop and cell phone can be used, if necessary, but **NO** game, **NO** noise, and **NOT** in the quizzes/exams! In any case, you are not allowed to disturb others in the classroom.

Other Course Policies:

- Information in this syllabus is subject to change with notice.
- Attendance of class and frequent check of email is expected. You are responsible for materials presented and announcements made in class or by email. This could

include changes to the syllabus, exam dates, etc. ***Three absences at class rolls will automatically exclude you from possible curving at final grades.***

- Make-up exams will only be given under extreme circumstances. The instructor reserves the right to reject make-up requests. There will be no make-up for unannounced quizzes (if any) under any circumstances.
- Be aware of the institution's policy on drops and incomplete.

University or Department Policies:

Prerequisite Proof (if the course has specific prerequisites listed above):

The Department of Computer Science has a policy that each instructor needs to verify the student transcript and ascertain that the student has the prerequisites. You can log on to My Sac State go to “Student Center” and select “Unofficial Transcripts” to print. You also can select and print “Transfer Credit Report” if you have transferred from another institution. You must submit your transcript for verification. Any student who has completed one or more prerequisites at another school must provide similar verification to the instructor. Any student who has not submitted their transcript for verification by the end of the second week will be dropped from the class.

Repeat Policy:

The department has a policy specifying that students may not repeat a computer science course more than once. Any student who wishes to repeat a course more than once (that is, take a course for a third time) must submit a petition requesting the permission to do so. Student records will be reviewed to determine whether a student is taking this course for three or more times. Any such student must return an approved petition to the instructor within the first two weeks of class. Any student who does not submit an approved petition will be dropped from the class. Petitions are available in the department office (RVR 3018) and require the signature of both the instructor and the department chair.

Drop Policy

If you plan to drop this course, please make sure you understand the following information.

- **There is no such thing as an “automatic drop”.** The instructor can drop you from the course, but this does not happen automatically. If you plan to drop the course, make sure to use MySacState.
- After the 2nd week, you cannot drop the course through MySacState. At this point, you must provide written verification of a compelling reason. Both the instructor and the Department Chair must approve.
- After the 4th week, you must fill out a “Petition to Drop after Deadline” form and collect all the necessary signatures. This must be turned into Admission and Records in Lassen Hall.

Students with Disabilities

If you have a disability and require accommodations, you need to provide disability documentation to SSWD (Services to Students with Disabilities), Lassen Hall 1008, (916) 278-6955. Please discuss your accommodation needs with me after class or in lab early in the semester.

Ethics/Academic Honesty

Any work submitted is a contractual obligation that the work is the student's and for which he/she could be quizzed in detail. Discussion among students in assignments and projects is part of the educational process and is encouraged. No discussion among students is allowed in any exams/quizzes. However, each student must make an effort to do his/her own work in all assignments and exams. No type of plagiarism will be tolerated except in the case of group work. In that case each student should indicate the part of the work, which was their major responsibility in their final joint submission. Nevertheless, I emphasize any work submitted is a contractual obligation that the work is the student's and for which he/she could be quizzed in detail. *The minimum penalty for even a single incident of cheating brought to the attention of the instructor in this course is automatic failure of the course; additional more severe penalties may also be applied. Note that cheating is grounds for dismissal from the University.*

Please refer to the Computer Science Dept. document entitled “Policy on Academic Integrity” (available online via the Computer Science department, www.ecs.csus.edu/csc home page) and to the University Policy Manual section on Academic Honesty (all available online via the instructor's home page. Please visit <http://www.csus.edu/admbus/umanual/UMA00150.htm>) for additional information. **IT IS THE RESPONSIBILITY OF EACH STUDENT TO BE FAMILIAR WITH, AND TO COMPLY WITH, THE POLICIES STATED IN THESE DOCUMENTS.** *In addition, unless otherwise stated, the use of the following devices during exams/quizzes is prohibited: cell phones, pagers, laptops, and PDAs.*

MSBA 206 TENTATIVE SCHEDULE SUBJECT TO CHANGE

Week	Topics and Major Content	Assignments	Reading Materials
5	Multiple Linear Regression <ul style="list-style-type: none"> - Introduction - Explanatory vs. Predictive Modeling 	Assignment 5	Textbook, Chapter 6
6	Multiple Linear Regression <ul style="list-style-type: none"> - Estimating the Regression Equation and Prediction - Multiple study cases - Variable Selection in Linear Regression 	Assignment 6	Textbook, Chapter 6
7	k-Nearest Neighbors (k-NN)-Application <ul style="list-style-type: none"> - The k-NN Classifier (Categorical Outcome) 	Assignment 7	Textbook, Chapter 7
8	k-Nearest Neighbors (k-NN)-Application <ul style="list-style-type: none"> - k-NN for a Numerical Outcome - Advantages and Shortcomings of k-NN Algorithms 	Assignment 8	Textbook, Chapter 7
9	Midterm Exam		
10	The Naive Bayes Classifier <ul style="list-style-type: none"> - Introduction - Applying the Full (Exact) Bayesian Classifier 	Assignment 9	Textbook, Chapter 8

	<p>The Naive Bayes Classifier</p> <ul style="list-style-type: none"> - Applying the Full (Exact) Bayesian Classifier - Advantages and Shortcomings of the Naive Bayes Classifier 	Assignment 10	Textbook, Chapter 8
11	<p>The Naive Bayes Classifier</p> <ul style="list-style-type: none"> - Applying the Full (Exact) Bayesian Classifier - Advantages and Shortcomings of the Naive Bayes Classifier 	Assignment 11	Textbook, Chapter 8
12	<p>Classification and Regression Trees</p> <ul style="list-style-type: none"> - Introduction - Classification Tree 		Textbook, Chapter 9
13	<p>Classification and Regression Trees</p> <ul style="list-style-type: none"> - Evaluating the Performance of a Classification Tree - Avoiding Overfitting - Multiple Classes Prediction 		Textbook, Chapter 9
14	<p>Classification and Regression Trees</p> <ul style="list-style-type: none"> - Regression Trees - Improving Prediction: Random Forests and Boosted Trees - Advantages and Weaknesses of a Tree 		Textbook, Chapter 9
15	Final Exam		