Course Syllabus: Data Mining and Text Analytics in Business

Dillard College of Business Administration
MIS 5613 Section Y70
Fall 2026 | August 24-December 4, 2026

Contact Information

Instructor: Jiaxi Luo
Office: Dillard Building 216
Office hours: Mondays & Wednesdays 3:30 - 5:30p Tuesdays 10:30 - 11:30a
Weekly lab: Wednesdays, 5:30-6:50 p.m.
Office phone: (940) 397-3268
E-mail: jiaxi.luo@msutexas.edu
Dr. Luo will try to respond to all emails within 24 hours on weekdays.

Course Description

Introduces data mining and text analytics methods for solving business problems. Students use Python to prepare and analyze structured and unstructured data; build and evaluate supervised and unsupervised models; and apply natural language processing methods for sentiment analysis, topic classification, summarization, and related tasks. Emphasis is placed on interpreting results, communicating actionable insights, and practicing responsible data use.

Textbook & Instructional Materials

Suggested readings (specific chapters and editions will be assigned in D2L):

- Machine Learning with Python Cookbook, by Chris Albon.
- Applied Text Analysis with Python, by Benjamin Bengfort, Rebecca Bilbro, and Tony Ojeda.
- Machine Learning Techniques for Text, by Nikos Tsourakis.
- Python Machine Learning, by Sebastian Raschka and Vahid Mirjalili.
- Python Machine Learning by Example, by Yuxi Liu.

Software Requirements

- Computer: A Windows or Mac laptop with a webcam and sufficient capacity to run PyCharm and required Python libraries. Chromebooks and tablets may not support the required software.
- Software: PyCharm Professional, available through a free student license, or another instructor-approved Python IDE, plus Python and assigned packages.
- Students need reliable internet access, regular access to D2L and MSU Texas e-mail, and the ability to create, run, save, and submit Python code and data files.

Course Objectives

General Learning Goals

- GLG1: Our students will exhibit the characteristics of leadership.
- Objective: Our graduates will have the capacity to lead effectively.
- GLG2: Our students will be able to conduct themselves professionally in a global business environment.
- Objective: Our graduates will be able to navigate through the differences among multicultural and international business environments.
- GLG3: Our students will produce creative responses to business situations.
- Objective: Our graduates will demonstrate the capability to critically analyze business situations and develop creative solutions to opportunities and problems.
- GLG4: Our students will integrate knowledge across business disciplines.
- Objective: Our graduates will demonstrate the capability to integrate knowledge across business disciplines.
- GLG5: Our students will communicate at a professional level.
- Objective: Our graduates will be able to communicate in a professional business manner.

Course-Specific Learning Goals

After completing this course, students should be able to:

- Explain fundamental data mining and text analytics concepts, including data preprocessing, feature extraction, classification, regression, clustering, and model evaluation.
- Use Python libraries such as pandas, NumPy, and scikit-learn to prepare, manipulate, analyze, and model business data.
- Use natural language processing libraries and methods to tokenize, normalize, represent, and analyze text data.
- Apply data mining and text analytics techniques to real-world business problems such as predictive modeling, sentiment analysis, topic classification, and summarization.
- Visualize and communicate analytical results clearly using appropriate Python tools, evidence, and business context.
- Complete hands-on analyses using real-world structured and unstructured data sets and collaborate effectively on a professional team project.
- Evaluate the ethical, legal, privacy, bias, and responsible-use implications of data mining, machine learning, and text analytics.

Student Handbook

Refer to: Student Handbook

Academic Misconduct Policy & Procedures

Academic Dishonesty: Cheating, collusion, and plagiarism (the act of using source material of other persons, either published or unpublished, without following the accepted techniques of crediting, or the submission for credit of work not the individual's to whom credit is given). Additional guidelines on procedures in these matters may be found in the Office of Student Conduct.
Office of Student Conduct

Moffett Library

Moffett Library provides resources and services to support student's studies and assignments, including books, peer-reviewed journals, databases, and multimedia materials accessible both on campus and remotely. The library offers media equipment checkout, reservable study rooms, and research assistance from librarians to help students effectively find, evaluate, and use information. Get started on this Moffett Library webpage to explore these resources and learn how to best utilize the library.

Grading

Table 1: Points allocated to each assignment

Assignments	**Points**
Exam 1	100
Exam 2	100
Final Exam	100
Team Project	150
Quizzes	100
Homework Assignments	100
Class Participation	50
Total Points	700

Table 2: Total points for final grade.

Grade	**Points**
A	630-700
B	560-629
C	490-559
D	420-489
F	Less than 420

Homework

Homework consists of individual Python programming assignments and is worth 100 points in total. From the data mining portion onward, most assignments include a subjective interpretation component. Unless otherwise stated in D2L, students must follow these submission requirements:

- **Format:** Submit standard .py files, with one .py file per question, and keep each file as runnable code. Interpretations may be included as comments in the corresponding .py file or collected in one Word document (.docx) that clearly references the related filenames.
- **Where:** Upload all required files to the corresponding D2L Assignment folder.
- **Due time:** Submit by 11:59 p.m. on the posted due date; the D2L timestamp is official.
- **Late work:** A deduction of 5% of the assignment's total points applies for each calendar day late. Work may be accepted for up to seven calendar days, with a maximum deduction of 35%; work more than seven days late is not accepted.
- **Before submitting:** Verify that every required file opens and runs correctly.

Quizzes

Quizzes are worth 100 points in total and are delivered regularly in D2L to reinforce each section's skills. Most questions are auto-graded, although occasional short answers may require manual review. Each quiz has a posted time limit and availability window and must be completed by 11:59 p.m. on the due date; late submissions are not accepted after the window closes.

Quizzes are individual work. Unless otherwise stated, students may use their notes and the internet for reference, but collaboration and answer-sharing are prohibited. The lowest non-bonus quiz score will be dropped. Scores and answer keys may be released after the availability window closes. Make-ups and accessibility accommodations must follow university policy and be arranged in advance.

Class Participation

Class Participation is worth 50 points and is based primarily on the quality, relevance, and timeliness of students' contributions to required D2L discussions, as well as preparation and meaningful engagement in Wednesday labs. Expectations and grading criteria will be explained in class and posted in D2L. One to four lab absences carry no point deduction. Five or more lab absences result in the loss of all 50 Class Participation points; lab absences affect no other grade category.

Team Project

Teams of three or four students will complete an integrated data mining and text analytics project. Each team must obtain its own data set from a public repository or collect one through ethical web crawling in accordance with the source site's policies. Ethical web crawling may earn extra credit as specified in D2L. The project must demonstrate a complete workflow with a strong text analytics component.

- Describe the data source, topic, size, structure, and attributes; document preprocessing with before-and-after examples; and present exploratory analysis with descriptive statistics and appropriate visualizations.
- Explain the selected classification, regression, or clustering task; train/test strategy; feature-selection decisions, if applicable; model choice; evaluation methods; and interpretation of results.
- Apply and justify an appropriate text analytics method, such as sentiment analysis, topic classification, or summarization, and communicate actionable findings, recommendations, and supporting visual evidence.
- Prepare a professional 10–15 minute recorded presentation that highlights the analytical story, important model choices, key findings, and proposed business solutions.
- Submit the Project Report in PDF or Word format, all runnable .py files and required data or supporting artifacts, the recorded presentation, and an individual peer evaluation in D2L. Teams that collect data must also submit their collection scripts and a brief note on policy compliance. Milestones, exact formatting requirements, and the grading rubric will be posted in D2L.

Exams

The course includes three individual online exams worth 100 points each: Exam 1 (Python programming), Exam 2 (data mining), and the Final Exam (natural language processing). Students may start at any time within the posted availability window. Once started, they have 5 hours to complete the exam, plus a 5-minute grace period to save and submit. Each exam contains 20 multiple-choice questions in D2L Quizzes and 5 coding questions completed outside the quiz. Exams are open-book, open-notes, and open-internet, but all work must be completed independently.

- **Multiple-choice questions:** Select and submit the answers directly in D2L Quizzes.
- **Coding questions:** Keep the quiz open and use D2L → Assignments → the corresponding Exam folder to download required files. Complete the work in an IDE and submit one .py file per question. Put short answers or interpretations in the corresponding .py comments or in one PDF, Word, or TXT file clearly keyed to the related filenames.
- **Use of AI:** AI may be used to assist with coding, but all subjective responses must be written by the student and grounded in the student's own results. When AI is used, include 2–4 sentences per coding question explaining what was verified or changed and why.
- **Required specifications:** Do not change required file or column names, and use only the model specified in each question.

Academic-integrity violations will be referred to the Office of Student Conduct and may result in course failure and additional disciplinary action. If the originality of a response is questioned, the student may be required to explain it verbally during office hours; the outcome may determine or adjust the exam's final score.

Late Work

Unless otherwise stated in D2L, late work will be penalized 5% of the assignment's total points for each calendar day after the deadline. Work submitted within seven calendar days may be accepted with a maximum deduction of 35%; work submitted more than seven calendar days late will not be accepted.

Make Up Work/Tests

Students with excused absences may make up missed examinations and course activities when supporting documentation is provided. Arrangements should be made in advance when possible. The instructor must be contacted no later than the day of the scheduled exam; otherwise, no make-up will be allowed. At the instructor's discretion, a deduction may be assessed for a late exam.
Excused absences include active military, police, or firefighter assignment; jury duty; university-authorized absences; and a medical emergency involving the student or an immediate family member. See the MSU Catalog for additional university policy information.

Experimental AI-Supported Oral Follow-Up

Following selected examinations, students may be required to complete an experimental AI-supported oral follow-up lasting approximately 10 minutes. The follow-up is considered part of the examination process and is intended to verify that students understand and can explain the work they submitted, including, as applicable, Excel formulas and workbooks, Tableau worksheets, visualizations and dashboards, Python code, analytical methods and decisions, results, observations, and written responses. Students should also be prepared to describe how any permitted AI assistance was used and how its output was reviewed, tested, or modified.

Students may be selected as part of a class-wide pilot, through random selection, or when clarification of submitted work is needed. The instructor may use the student’s oral responses to confirm or adjust the score for the related question(s) or portion(s) of the examination. A student’s inability to explain, interpret, reproduce, or verify essential elements of submitted work may result in a reduction of the corresponding examination score. Evidence of unauthorized assistance, prohibited collaboration, or misrepresentation of submitted work may also be addressed under the Academic Misconduct Policy.

Detailed procedures, scheduling information, evaluation criteria, technology requirements, and practice opportunities will be provided in D2L before the first oral follow-up. Approved accommodations and make-up arrangements will be provided in accordance with University policy.

Important Dates

- Last day for term schedule changes: August 24-27, 2026
- Deadline to file for graduation: September 21, 2026
- Last Day to drop with a grade of "W:" November 23, 2026, at 4:00 p.m.
- Refer to: Drops, Withdrawals & Void

Desire-to-Learn (D2L)

Extensive use of the MSU D2L program is a part of this course. Each student is expected to be familiar with this program as it provides a primary source of communication regarding assignments, examination materials, and general course information. You can log into D2L through the MSU Homepage. If you experience difficulties, please contact the technicians listed for the program or contact your instructor.

Attendance

Students are expected to attend all meetings of the classes in which they are enrolled. Although students generally are graded on intellectual effort and performance rather than attendance, absences may lower a student's grade when attendance and participation are essential. An instructor with an attendance policy must keep daily records and give a verbal or written warning before dropping a student from the course. A student with excessive absences may be dropped in accordance with General University Policies.

Online Computer Requirements

Taking an online class requires you to have access to a computer (with Internet access) to complete and upload your assignments. It is your responsibility to have (or have access to) a working computer in this class. ***Assignments and tests are due by the due date, and personal computer technical difficulties will not be considered reason for the instructor to allow students extra time to submit assignments, tests, or discussion postings.***

Computers are available on campus in various areas of the buildings as well as the Academic Success Center. ***Your computer being down is not an excuse for missing a deadline!!** There are many places to access your class! Our online classes can be accessed from any computer in the world which is connected to the internet. Contact your instructor immediately upon having computer trouble If you have technical difficulties in the course, there is also a student helpdesk available to you. The college cannot work directly on student computers due to both liability and resource limitations however they are able to help you get connected to our online services. For help, log into D2L**.**

Recording Classes

Recording classes is generally prohibited without the professor's explicit permission. Such recordings, when permitted, are for personal use only and may not be uploaded to the internet or otherwise shared, transmitted, or published without the professor's prior consent. Some faculty may not permit recordings of their classes. When it is feasible and appropriate, students may record classes to accommodate disabilities under the Americans With Disabilities Act, with prior approval of Disability Support Services. Such recordings may not be uploaded to the internet or otherwise shared, transmitted, or published.

Instructor Class Policies

AI-Tool Policy

Students may use AI tools such as ChatGPT or GitHub Copilot for the following limited purposes:

- English writing: Grammar and syntax improvement.
- Coding support: Explaining Python, data mining, or NLP concepts; suggesting troubleshooting steps; and improving code quality.
- Project planning: Brainstorming research questions, preprocessing options, model comparisons, text-analysis approaches, and visualization ideas.

The following requirements apply to every use of AI tools:

- Original analysis and code: Students must understand, test, and be able to explain submitted code, model choices, interpretations, and conclusions.
- Academic integrity: Students must disclose permitted AI assistance, verify all output, cite it when required, avoid over-reliance, and never upload protected or confidential data.

AI tools may be used during exams only as explicitly authorized in the Exams section and the corresponding D2L instructions. Use AI responsibly and prioritize genuine understanding, accurate work, and original thinking.

Change of Schedule

The instructor may modify the course schedule, procedures, assignments, or deadlines when necessary. Changes will be announced in D2L. Students are responsible for following the most recently posted schedule and announcements.

Refund and Repayment Policy

A student who withdraws or is administratively withdrawn from Midwestern State University (MSU) may be eligible to receive a refund for all or a portion of the tuition, fees and room/board charges that were paid to MSU for the semester. HOWEVER, if the student received financial aid (federal/state/institutional grants, loans and/or scholarships), all or a portion of the refund may be returned to the financial aid programs. As described below, two formulas (federal and state) exists in determining the amount of the refund. (Examples of each refund calculation will be made available upon request).

Services for Students with Disabilities

In accordance with Section 504 of the Federal Rehabilitation Act of 1973 and the Americans with Disabilities Act of 1990, Midwestern State University endeavors to make reasonable accommodations to ensure equal opportunity for qualified persons with disabilities to participate in all educational, social, and recreational programs and activities. After notification of acceptance, students requiring accommodations should make application for such assistance through Disability Support Services, located in the Student Wellness Center, (940) 397-4140. Current documentation of a disability will be required in order to provide appropriate services, and each request will be individually reviewed. For more details, please go to Disability Support Services.

Emergency Aid Resources for Basic Needs

Unexpected life circumstances can affect anyone. If you are experiencing challenges meeting basic needs—such as housing, utilities, food, transportation, healthcare, technology, textbooks or other life necessities, resources may be available to help. Visit the Student Emergency Support and Resources webpage to learn more, and explore emergency support opportunities through MSU and how to access them. For additional information, please contact the Office of Student Affairs:
Clark Student Center #108, (940) 397-7500, student.affairs@msutexas.edu.

College Policies

Campus Carry Rules/Policies

Refer to: Campus Carry Rules and Policies

Smoking/Tobacco Policy

MSU Texas is a tobacco-free campus. Smoking and the use of smokeless tobacco, e-cigarettes, and other tobacco-derived products are prohibited in all University-owned or leased indoor and outdoor areas, University vehicles, and private vehicles parked on University property. The policy applies to everyone, with no exceptions.

Alcohol and Drug Policy

To comply with the Drug Free Schools and Communities Act of 1989 and subsequent amendments, students and employees of Midwestern State are informed that strictly enforced policies are in place which prohibits the unlawful possession, use or distribution of any illicit drugs, including alcohol, on university property or as part of any university-sponsored activity. Students and employees are also subject to all applicable legal sanctions under local, state and federal law for any offenses involving illicit drugs on University property or at University-sponsored activities.

Campus Carry

Effective August 1, 2016, the Campus Carry law (Senate Bill 11) allows those licensed individuals to carry a concealed handgun in buildings on public university campuses, except in locations the University establishes has prohibited. The new Constitutional Carry law does not change this process. Concealed carry still requires a License to Carry permit, and openly carrying handguns is not allowed on college campuses. For more information, visit Campus Carry.

Active Shooter

The safety and security of our campus is the responsibility of everyone in our community. Each of us has an obligation to be prepared to appropriately respond to threats to our campus, such as an active aggressor. Please review the information provided by MSU Police Department regarding the options and strategies we can all use to stay safe during difficult situations. For more information, visit MSUReady – Active Shooter. Students are encouraged to watch the video entitled "*Run. Hide. Fight.*" which may be electronically accessed via the University police department's webpage: *"Run. Hide. Fight."*

Grade Appeal Process

Students who wish to appeal a grade should consult the Midwestern State University MSU Catalog

Weather Procedure

In the event of inclement weather, in-person lab meetings may be canceled. Alternative online tasks, recorded instruction, or video-based activities may be assigned through D2L to support continued learning.

*__Notice:__ Changes in the course syllabus, procedure, assignments, and schedule may be made at the discretion of the instructor.

Course Schedule:

Keep this syllabus as a reference. Students are responsible for its contents and for announced changes. The schedule below follows the Fall 2026 weekly cadence and may be adjusted by the instructor.

Week	Date	Format	Topic
1	08/26/2026 – 08/31/2026	Online	Python Basics I: Variables, Data Types, and Input/Output
1	08/26/2026	Lab	Practice: Python Installation, Course Introduction, and PyCharm Setup
2	08/31/2026 – 09/07/2026	Online	Python Basics II: Functions and Control Flow
2	09/02/2026	Lab	Practice: Writing Functions and Using Loops/Conditionals
3	09/07/2026 – 09/14/2026	Online	Python Collections
3	09/09/2026	Lab	Practice: Manipulating Lists, Tuples, and Dictionaries
4	09/14/2026 – 09/21/2026	Online	Python Modules and Visualization
4	09/16/2026	Lab	Practice: Using Math/Statistics Modules and Creating Basic Charts with Matplotlib
5	**09/25/2026 – 09/26/2026**	**Exam**	**Exam I: Python**
5	09/21/2026 – 09/28/2026	Online	Data Mining Concepts, Data Preparation, and Model Evaluation
5	09/23/2026	Lab	Practice: Exploring Business Datasets with Pandas
6	09/28/2026 – 10/05/2026	Online	Classification Models I: Regression and Decision Trees
6	09/30/2026	Lab	Practice: Regression and Decision Trees with scikit-learn
7	10/05/2026 – 10/12/2026	Online	Classification Models II: Naive Bayes, SVM, Neural Networks, and Clustering
7	10/07/2026	Lab	Practice: Naive Bayes, SVM, Neural Networks, and Clustering with scikit-learn
8	10/12/2026 – 10/19/2026	Online	Data Collection: APIs and Web Scraping
8	10/14/2026	Lab	Practice: Collecting Data with APIs and Web Scraping
9	**10/23/2026 – 10/24/2026**	**Exam**	**Exam II: Data Mining**
9	10/19/2026 – 10/26/2026	Online	Natural Language Processing Introduction
9	10/21/2026	Lab	NLP Introduction
10	10/26/2026 – 11/02/2026	Online	NLP Concepts: Text Representation and Preprocessing
10	10/28/2026	Lab	NLP Concepts
11	11/02/2026 – 11/09/2026	Online	NLP Applications: Sentiment Analysis, Classification, and Summarization
11	11/04/2026	Lab	NLP Applications
12	11/09/2026 – 11/16/2026	Online	NLP and Large Language Models
12	11/11/2026	Lab	NLP Large Language Model
13	11/16/2026 – 11/23/2026	Online	NLP Case Studies in Business
13	11/18/2026	Lab	NLP Case Study
14	11/23/2026 – 11/30/2026	Online	NLP Visualization and Evaluation
14	**11/25/2026**	**Holiday**	**Thanksgiving Break — No Lab**
15	11/30/2026 – 12/04/2026	Online	Project Showcase and API Integration
15	12/02/2026	Asynchronous	Final Exam Review — No In-Person Lab Meeting
15	**12/06/2026 – 12/07/2026**	**Exam**	**Final Exam: Natural Language Processing**