UNIVERSITY OF CALIFORNIA, DAVIS

GRADUATE SCHOOL OF MANAGEMENT

AI AND BUSINESS INNOVATION:
A JOURNEY FROM LINEAR REGRESSION TO LARGE LANGUAGE MODELS

(MGV-490A, FALL 2024)

SYLLABUS (09/22/2024)

Instructor: Jörn Boehnke, Assistant Professor

Office: 3404, Gallagher Hall, Graduate School of Management

E-Mail: jb@ucdavis.edu

In-person: University of Nevada, Reno at Lake Tahoe – Incline Village
999 Tahoe Blvd, Incline Village, NV 89451

Friday (10/4) 1pm-4pm
Saturday (10/5) 9am-12pm & 1pm-4pm
Sunday (10/6) 9am-12pm

Office Hours: By appointment via Zoom.

Please be encouraged to reach out to me via email at any point!

Heads-up: This course has no prerequisites, but it will be rigorous and demanding. We will explore data analytics and AI from a strategic, high-level perspective. Your effort will directly impact what you gain, and it will be highly rewarding. Join me to enhance your data-savviness and learn to make informed, data-driven business decisions.

Objectives: After taking this course, you will be equipped to explore and make sense of large, complex business datasets, gain insights into creating and using powerful predictive systems for business applications, and learn how to interpret the results of these models.

Before diving into the main topics, we will start with a very, very gentle introduction to Python programming for Data Analytics. This asynchronous content is designed for business professionals with little to no programming experience. You will learn basic Python syntax and how to use essential libraries for data analysis. By the end of it, you will have the foundational skills needed to implement and understand the computational techniques covered in the rest of the course.

The list of topics is:

Asynchronous
A *very, very gentle* introduction to Python programming for Data Analytics (recorded content + exercises ~8h)

Read (or listen to its audiobook) "Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking"
- the first 3 chapters are due *prior* to our in-person section
- the remaining 11 chapters (4-14) are due *after* we meet in person

In-person
Introduction
Regression
Treatment Effect
Model Selection
Classification
Clustering
Trees
Neural Networks
Natural Language Processing
Large Language Models

We will cover both basic concepts and practical skills, including techniques for analyzing distributed data. A strong emphasis is placed on working with actual business datasets and developing methods tailored to specific applications. Examples will include consumer data mining, internet and social media tracking, sports analytics, and text mining.

Asynchronous Learning: The first portion of this course will be offered asynchronously, allowing you to learn at your own pace. It includes a gentle introduction to Python programming for Data Analytics (~8 hours of recorded content and exercises) and an assigned reading/listening of "Data Science for Business." This will provide you with the necessary foundational skills before moving on to the more advanced, in-person sessions.

TA: Tiruo Yan (tyyan@ucdavis.edu)

Regular TA office hours (Zoom) will be offered in the weeks leading up to our in-person session to optimally support the asynchronous material.

Please also feel free to reach out to your TA via email at any point.

Evaluation:
10% Assignments before in-person section
10% Mini exam Friday
10% Mini exam Saturday
10% Mini exam Sunday
20% Final group project
20% Final exam
20% Class participation

Late assignments, exams, and projects will not be accepted.

Clerical scoring errors will be corrected without hassle, but for other re-grades you must hand back the work and submit an email request; the entire paper will be subject to re-grading.

Groups: You will be assigned to a group at the beginning of the quarter to discuss assignments and work on the final project. *Please note that group work will be subject to a peer assessment at the end. One-half of your total group work grade (assignments and final project) will be determined by the score on the peer evaluations.* I.e., individuals who do not contribute their fair share to the group (as determined by the group) will be penalized.

Class Participation: I hope you actively participate in class. Meaningful class interactions add greatly to the learning experience, and I consider class participation to be an important component of this course and its success. Accordingly, active and meaningful participation in class and on Piazza is part of your grade evaluation. Please be prepared to be cold-called in class.

Assignments: There will be three homework assignments before we meet in-person. The assignments will cover the asynchronous material of the class (online lectures + content of the book). Assignments are submitted through Canvas and should have a clear and concise presentation.

All assignments are individual assignments. Assignments *may be* discussed among your group members but *must be* submitted individually. While your code submission may be similar to that of your group's members, any write-up, code comments, etc. must be your own. Simply copying and pasting the write-up and code comments from other members in your group will yield point deductions. I want each group member to fully understand the code and analytics required in assignments.

Assignments are not to be discussed outside your group. Any violation of this policy will be considered plagiarism and penalized as such.

My two cents: I strongly encourage you all to work together in your group. Meet with your group regularly to discuss and learn from each other. It's only the code annotations / interpretation of the results that I want each of you to do individually.

Classroom Etiquette / Extra Assignments: Classes begin on time. You are required to attend every class unless extenuating circumstances arise and make it difficult to do so. If you are unable to attend a class, you are required to inform me about the extenuating situation by email *at least one hour* before the start of class.

All cell phones must be muted before the start of class. The computer should only be used to take class notes and code / run in-class exercises. All other programs should be shut down before the start of class. Any messaging, web surfing, e-mailing, etc. disturbs the class and is a breach of classroom etiquette. If it comes to my attention, any such behavior results in *an extra assignment*.

Exams: There will be three mini exams as well as a comprehensive final exam. The mini exams will be conducted in the first 5-10 minutes of lecture each day (in-person).

Mini exam 1 will cover the concepts learned in the asynchronous section of this course. Mini exams 2 and 3 will cover the concepts learned during the in-person section.

The final exam will cover everything we have learned asynchronously as well as in-person. **The final exam will be conducted online approximately two weeks after our in-person section ends.**

Project: In the final project you will study a business problem of your choice. The goal of the project is to describe and analyze data for business use. Specifically, your analysis ought to add insights that help the business problem at hand. For this project, think of yourselves as consultants who want to support some company's business decisions. These decisions should be based on insights gained from analyzing the data and it is your task to provide this analysis.

Planning should begin during our in-person section. Successful projects will implement a variety of machine learning concepts we covered in class. All implementations must be documented, and all code must be attached to your submission. Any number of primary and secondary data sources can be used for this project. Please always adhere to the terms and conditions pertaining to the websites and data you use.

Typically, project reports will involve the following components: (1) Title Page, (2) Executive Summary, (3) Background, Context, and Domain Knowledge: business (scenario) in mind, type of industry, products / services, competition, (4) Discussion of how the firm / industry traditionally attempts to solve this problem, the specific strategy, and how it aligns with the business model, (5.1) Analyses, (5.2) More Analyses, (5.3) Even More Analyses, (6) Recommendations and Business Value provided, and finally, (7) Summary and Conclusions.
(5) and (6) will showcase what you learned in class and how well you can apply it. Please make sure that these components form the centerpiece of your project report.

The final project report cannot exceed 10 pages (not including the title page, code attachments, references, and appendices). Any necessary tables, figures, visualizations, and text must be contained within the 10 pages. I urge you to ensure that the written report is direct, insightful, and specific to the problem at hand. The report should adhere by the following formatting guidelines: text no smaller than 11-point font, 1-inch

margins on all pages, and all text should be double-spaced. The report should also contain an executive summary, which counts towards the page limit. **All submissions are due two weeks after our in-person section ends.** No exceptions.

Software: In the evolving landscape of business analytics, Python has emerged as the predominant programming language, widely recognized for its versatility and effectiveness. Reflecting this industry trend, this course will exclusively use Python for all computing tasks. This focus will allow us to dive deeply into its application in business analytics, providing you with relevant and practical skills.

Python's extensive use in data analysis and its robust libraries make it a great tool for business analysts. During this course, all homework and class examples will be in Python, ensuring a consistent and focused learning experience. While a basic understanding of Python is expected at the start of the class, be prepared for a steep learning curve if you are new to it.

Python does not need to be installed from scratch on your machine. We will utilize online environments like Google Colab to execute all Python code developed in this course, ensuring accessibility and ease of use for everyone.

This course is not designed to make you a software developer but rather to provide a gentle introduction to Python, specifically for use in business analytics. Our aim is to equip you with the basic Python skills needed to fluently interact with software developers and remove communication barriers between business, data analytics, and software implementation. The foundational knowledge we cover will enable you to effectively collaborate with technical teams and leverage data-driven insights in your business decisions.

Rules: All participants in the course, instructor and students, are expected to follow the UC Davis Principles of Community, which includes affirmation of the right of freedom of expression, and rejection of discrimination. The right to express points-of-view without fear of retaliation or censorship is a cornerstone of academic freedom. A diversity of opinions with respectful disagreement and informed debate enriches learning. However, in this course, any expression or disagreement should adhere to the obligations we have toward each other to build and maintain a climate of mutual respect and caring.

You are expected to take UC Davis's Code of Academic Conduct as seriously as we do. You were given this code of conduct with explicit explanations of violations (e.g. plagiarism, cheating, unauthorized collaboration, etc.) and your responsibilities in regard to them during orientation, and you signed a statement affirming that you understand it. Academic conduct violations will not be tolerated, and your instructor will not hesitate to turn violators over to Student Judicial Affairs. If you are uncertain about what constitutes an academic conduct violation, please refer to the code linked above, contact your instructor, or refer to the Office of Student Judicial Affairs.

Q&A: We are using Piazza for class discussion (accessible through Canvas). The system is highly catered to getting you help fast and efficiently from classmates, the TA, and myself. Rather than emailing your questions, try to post on Piazza. If you have any problems, you can email me or team@piazza.com.

Please feel encouraged to answer your classmate's questions; it's a huge help to us, and even if you are wrong everyone learns (we check the answers and clear up confusion). While you can post anonymously, we encourage you to take credit for your questions and answers.

Texts: For this course, you are required to read "Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking" by Foster Provost and Tom Fawcett. This book provides a comprehensive introduction to the fundamental principles and concepts of data science, tailored for business professionals. It offers valuable insights into data mining, predictive modeling, and data-analytic thinking, making it an essential resource for understanding how data-driven decisions can transform business strategies. Notably, this is one of only a few technical books that are available as audiobook, making it convenient to consume on the go.

"Data Science for Business" is the only required textbook; all other required materials will be available on the class page.

For those interested in further expanding their knowledge, the following books are recommended but not required: *An Introduction to Statistical Learning: with Applications in Python*, by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, and Jonathan Taylor. This is a great intermediate-level book that provides a solid foundation in statistical learning techniques, complete with practical examples and applications in Python.

A great advanced text is *The Elements of Statistical Learning* by Trevor Hastie, Robert Tibshirani, and Jerome Friedman, but it requires some mathematical sophistication and goes beyond the material we will be covering. Also good, but still advanced, are *Pattern Recognition & Machine Learning* by Christopher M. Bishop. This book introduces the material from a more computer engineering, rather than statistical, perspective.

Accommodation: UC Davis is committed to educational equity in the academic setting, and in serving a diverse student body. All students who are interested in learning about how disabilities are accommodated can visit the Student Disability Center (SDC). If you are a student who requires academic accommodations, please contact the SDC directly at sdc@ucdavis.edu or 530-752-3184. If you receive an SDC Letter of Accommodation, submit it to your instructor for each course as soon as possible, at least within the first two weeks of a course.

Safety and Emergency Preparedness: UC Davis has many resources to help in case of emergency or crisis. While reviewing campus Emergency Information, you may want to register for UC Davis Warn Me and Aggie Alert, which will give you timely information and instructions about emergencies and situations on campus that affect your safety.

If there is an emergency in the classroom or in non-Davis locations, follow the instructions of your instructor.

Disclaimer: Unexpected events might require elements of this syllabus to change. Your instructor will keep you informed of any changes.