THE OHIO STATE UNIVERSITY

The Max M. Fisher College of Business
Department of Operations and Business Analytics
BUSOBA 7332- Predictive Analytics
Term I, Spring Semester 2022

Instructor:

Name: Dr. Ismael Talke
Department: Operations and Business Analytics
Office Location: Fisher Hall 334/[Click Zoom Link]
Phone Number: 614-688-2452
Email: talke.1@osu.edu
Office Hours (Virtual Via Zoom. Click Zoom Link): TH & Fri 6:30:-7:30PM; Other time by appt.

Course Materials / Software:

Textbooks (digital copies of both texts are available online through the OSU library):

- **Primary text (APM):** Max Kuhn and Kjell Johnson, Applied Predictive Modeling, Springer (2013). Companion APM Book Website and corresponding AppliedPredictiveModeling R package.
- **Supplementary text (ISLR):** Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, An Introduction to Statistical Learning: With Applications in R, Springer (2013). Companion ISLR Book Website and corresponding ISLR R package.
- Videos, lecture notes, tutorials, data sets, and various supplementary materials which will made available through the course website on Carmen.
- **Software**: R Studio, R Statistical Package and related tools

Prerequisites: Enrollment in Fisher's Specialized Master's Degree Program in Business Analytics (SMB-A), BUSMGT 7250 or permission of the instructor.

Course Overview: Business analytics is a process of exploring and transforming data to gain insight and using it to predict outcomes, trends and patterns. This is an applied course and students will learn, through hands on approach the different methods of predictive analytics concepts and techniques in solving practical real-world business problems using R language and relevant R-based packages. As an applied course, the emphasis will be less on the technical/mathematical underpinning of each method and more on the key intuition behind it, when and how to properly use it, how to tune it for optimal performance, and how to interpret and evaluate results. The techniques covered in this course fall into two major categories: (1) supervised learning techniques, including regression and classification, decision trees, and Neural

networks; and (2) unsupervised learning techniques, cluster analysis and Market Basket Analysis. These methods have broad applications across all functional areas within business organizations; they have been successfully applied in areas as diverse as churn analysis, direct marketing, fraud detection, consumer retention, click-stream analysis, risk management, credit scoring and others.

At the end of this course, students should be able to

1. Organize and prepare data to use for predictive modeling.
2. Understand and describe the different types of modeling problems and models.
3. Develop appropriate prediction models using R, assess the prediction power of the model, and interpret results to solve business problems.
4. Effectively communicate the results and present report.

How This Course Works:

Mode of delivery: Consistent with the delivery model for courses in the SMB-A program, this course is of a distance-hybrid structure. Material is presented asynchronously online on a weekly cadence. Reading will be assigned from selected eBook chapters, tutorials and notes with other online supplemental materials. Some materials will be covered only in the readings, some others will be covered only in video which may depart from the text in content and order.
On 3 Saturdays in the semester (**January 22nd, February 5th and 19th, 1:30pm-5:00pm**), students will be required to attend or stream (in person or synchronously) a 3.5-hour session in which activities and discussion will occur.

Pace of online activities:
Students are expected to keep pace with weekly deadlines but may schedule their efforts freely within that time frame. Because most of the formal instruction will occur asynchronously online, a higher level of self-discipline is required in order to successfully complete this course. You cannot afford to get behind. It is therefore strongly recommended that you set up a personal study schedule that specifically earmarks time when you will regularly work on the material for this course. If you do not stay up with the course schedule provided in this syllabus you will not successfully complete this course.

Credit hours and work expectations:
This is a **3-credit-hour course delivered in 7 weeks (1/2 semester)**. According to Ohio State policy, students should expect around an average of 3 hrs per week of time spent on direct instruction. Specifically rule 3335-8-24, credits at Ohio State work on a 1 to 3 ratio: every 1 semester credit hour assigned to the class equates to total of 3 hours of work per week (1 hour of instruction and 2 additional study hours per week). Because this course is delivered over half a semester (7-weeks), this expectation doubles.

Attendance and participation requirements:
Because this is a hybrid course, your attendance is based on your online activity and participation. The following is a summary of everyone's expected participation:

• **Participating in online activities for attendance**: AT LEAST TWICE PER WEEK, You are expected to log in and study assigned materials on Carmen every week. (During most weeks you will probably log in more than that.) If you have a situation that might cause you to miss an entire week of class, discuss it with me as soon as possible.

• **On-campus class sessions**: 3 MEETINGS (Saturdays), you are also to attend in-person class sessions (in-person or streaming).

• **Virtual Office hour and live sessions:** via Zoom and participating in discussion forums: OPTIONAL As part of your participation, each week we suggest you participate in class discussion on the discussion forum on Carmen.

Course technology:

For help with your password, university email, Carmen, or any other technology issues, questions, or requests, contact the Ohio State IT Service Desk. Standard support hours are available at ocio.osu.edu/help/hours, and support for urgent issues is available 24/7.

- **Self-Service and Chat support:** ocio.osu.edu/help
- **Phone:** 614-688-4357(HELP)
- **Email:** servicedesk@osu.edu
- **TDD:** 614-688-8743

Baseline technical skills for online courses

- Basic computer and web-browsing skills
- Navigating Carmen: for questions about specific functionality, see the Canvas Student Guide.

Required Technology skills specific to this course

- CarmenZoom virtrual meetings
- Basic Excel literacy

Required equipment

- Computer: current Mac (OS X) or PC (Windows 7+) with high-speed internet connection
- Webcam: built-in or external webcam, fully installed and tested
- Microphone: built-in laptop or tablet mic or external microphone
- Other: a mobile device (smartphone or tablet) or landline to use for BuckeyePass authentication

Required software

- Microsoft Office 365: All Ohio State students are now eligible for free Microsoft Office 365 ProPlus through Microsoft's Student Advantage program. Full instructions for downloading and installation can be found at go.osu.edu/office365help.

Carmen access: You will need to use BuckeyePass multi-factor authentication to access your courses in Carmen. To ensure that you are able to connect to Carmen at all times, it is recommended that you take the following steps:

- Register multiple devices in case something happens to your primary device. Visit the BuckeyePass - Adding a Device help article for step-by-step instructions.
- Request passcodes to keep as a backup authentication option. When you see the Duo login screen on your computer, click **Enter a Passcode** and then click the **Text me new codes** button that appears. This will text you ten passcodes good for 365 days that can each be used once.
- Download the Duo Mobile application to all of your registered devices for the ability to generate one-time codes in the event that you lose cell, data, or Wi-Fi service.

If none of these options will meet the needs of your situation, you can contact the IT Service Desk at 614-688-4357 (HELP) and IT support staff will work out a solution with you.

Grading and Evaluation:

Graded assignments may come in three forms, and students should note the expectations for each in the descriptions of our class assignments below.

- **(N) Independent Work ( ):** Strictly non-collaborative, original-individual work. You may discuss this assignment only with your instructor. Discussions with other individuals, either in person or electronically, are strictly prohibited.
- **(RC) Required Collaboration ( ):** An explicit expectation for collaboration among students either in-class or outside (i.e. group work).
- **(OC) Optional Collaboration ( ):** Students are permitted, but not required, to discuss the assignment or ideas with each other. However, all submitted work must be one's original and individual creation.

Assignment Name	Points / Weight	Assignment Type
Homework	100 Pts / 40%	
Course Project	50 Pts / 20%	
Activity Participation	25 Pts / 10%	
Final Exam	75 Pts / 30%	
TOTAL	250 Pts / 100%	

Final grades will be based on overall class performance.

Course Assignments:

Homework: In addition to the reading requirements, there will be 5-6 weekly assignments. They are designed to reinforce and test your understanding of the topics covered. Unless otherwise

indicated, all assignments are due **at 11:59PM on Monday** by the following week. Assignments are to be submitted on or before the designated due date and time. Late work will not be accepted for credit. While the homework assignments may vary in length and/or difficulty, each will be graded out of a possible 20 points. Your lowest homework score will not count toward your total assignment score in the course, so you can miss one homework without it counting toward your course grade.

A limited amount of cooperation among students on homework and lab assignments is permitted. You may discuss with classmates general solution strategies; however, everyone should independently do and turn in his/her own work. If you use any references, including solutions to similar problems found online or prepared by other students, you must cite and credit your sources.

Team-Based Term Project

Students will have the opportunity to put to work the tools and knowledge gained in this course, to sharpen their skills and acquire hands-on experience with real-world data sets through a term project. The course project will be carried out in teams of 3-4 students. Teams need to be formed by **January 27th**. More information about the course project will follow later.

Final Exam

There will be one final exam. The examination is designed to assess each student's (a) command of factual knowledge and concepts from the course; and (b) his or her ability to integrate and generalize these concepts and principles and apply them to new situations. The final exam must be taken at the scheduled time; a make-up exams will only be given for truly special and compelling cases, in accordance with University guidelines. You may reference any and all course resources (book/notes, and R codes) during the final exam. You may **NOT** communicate with any other persons (physically, virtually, electronically, or otherwise) during the final exam. The final exam will **open Monday, February 28th at 6:30pm** and **it will close Wednesday, March 2nd at 11:59pm.** You have more than 2 days to complete the test.

Activity participation

This will be based on your online activity and participation in Carmen discussion forum as well as in the in person or streaming session.

Late Assignment Submissions: As all assignments are virtually submitted, no late homework or case assignments will be accepted without prior approval. Due to the current situation, a level of understanding will be accorded if you contact the professor before the assignment is due.

Instructor Feedback and Response Expectations:

Email Communication: The subject line **must include BUSOBA7332**. Expect response within 1 business day. Most responses will be significantly quicker.

Safety and health requirements:

All teaching staff and students are required to comply with and stay up to date on all University safety and health guidance, which includes wearing a face mask in any indoor space and maintaining a safe physical distance at all times. Non-compliance will be warned first and disciplinary actions will be taken for repeated offenses.

Academic integrity:

Academic integrity is essential to maintaining an environment that fosters excellence in teaching, research, and other educational and scholarly activities. Thus, The Ohio State University and the Committee on Academic Misconduct (COAM) expect that all students have read and understand the University's Code of Student Conduct, and that all students will complete all academic and scholarly assignments with fairness and honesty. Students must recognize that failure to follow the rules and guidelines established in the University's Code of Student Conduct (https://trustees.osu.edu/bylaws-and-rules/code) and this syllabus may constitute Academic Misconduct (https://oaa.osu.edu/academic-integrity-and-misconduct)

The Ohio State University's Code of Student Conduct (Section 3335-23-04) defines academic misconduct as: Any activity that tends to compromise the academic integrity of the University, or subvert the educational process. Examples of academic misconduct include (but are not limited to) plagiarism, collusion (unauthorized collaboration), copying the work of another student, and possession of unauthorized materials during an examination. Ignorance of the University's Code of Student Conduct is never considered an excuse for academic misconduct, so I recommend that you review the Code of Student Conduct and, specifically, the sections dealing with academic misconduct.

If I suspect that a student has committed academic misconduct in this course, I am obligated by University Rules to report my suspicions to the Committee on Academic Misconduct. If COAM determines that you have violated the University's Code of Student Conduct (i.e., committed academic misconduct), the sanctions for the misconduct could include a failing grade in this course and suspension or dismissal from the University.
If you have any questions about the above policy or what constitutes academic misconduct in this course, please contact me.

It is also expected that each student will behave in a manner that is consistent with the Fisher Honor Statement, which reads as follows:

As a member of the Fisher College of Business Community, I am personally committed to the highest standards of behavior. Honesty and integrity are the foundations from which I will measure my actions. I will hold myself accountable to adhere to these standards. As a future leader in the community and business environment, I pledge to live by these principles and celebrate those who share these ideals.

Disability Services:

The university strives to make all learning experiences as accessible as possible. In light of the current pandemic, students seeking to request COVID-related accommodations may do so through the university's request process, managed by Student Life Disability Services. If you anticipate or experience academic barriers based on your disability (including mental health, chronic, or temporary medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion. **SLDS contact information:** slds@osu.edu; 614-292-3307; slds.osu.edu; 098 Baker Hall, 113 W. 12th Avenue.

Grievances and Solving Problems:

According to University Policies, if you have a problem with this class, you should seek to resolve the grievance concerning a grade or academic practice by speaking first with the instructor or professor. Then, if necessary, take your case to the department chairperson, associate dean for programs in the college, and to the provost, in that order. Specific procedures are outlined in Faculty Rule 3335-7-23. Grievances against graduate, research, and teaching assistants should be submitted first to the supervising instructor, then to the chairperson of the assistant's department

Copyright:

Tentative Course Schedule:

The following schedule gives the general plan for the course; changes may be made at my discretion but are designed to optimize the quality and flow of the content. The course web site gives the dynamic picture and is an integral part of the class; please make sure to check it on a regular basis.

Week	Topic(s)	Activities: Reading (R), Video (V) and Other (O)	Due Date 11:59pm	Objectives
1.	**Course introduction** ✓ Syllabus overview ✓ Intro. To predictive analytics. ✓ Basic foundations ✓ Review of regression models. ✓ R refresher	(R) APM Ch. 1, 2 and 3. (R) ISLR Ch. 1, 2.1, 2.3, 3.1-3.3 (V) from instructor		- To understand the foundations concepts of predictive analytics. -organizing and preparing data -Refresh Regression - To introduce R software for predictive analytics
2.	**Broad concept of model building** ✓ Data Pre-processing ✓ Data Partitioning ✓ Performance measures ✓ Overfitting problem ✓ Model Tuning using regularization methods ○ Ridge, ○ LASSO and ○ Elastic Net	(R)APM Ch. 4, 5 and 6.4 and 6.5, 6.61 (R) ISLR Ch. 2.2, Ch. 5 and Ch. 6.2 (V) from instructor	HW1 1/18	-To work with different types of data types, and how to prep-process the data for analysis, model performance, address over fitting problem using resampling and model tuning (regularization methods).
3.	**Classification Models** ✓ Logistic Regression ✓ Classification metrics ✓ Assessing model, ✓ Multinomial logistic regression	(R) APM Ch. 12 (R) ISLR Ch 4 (V) from instructor	HW2 1/24	-To learn and understand how to fit predictive models when the outcome is binary or has more than 2 categories using the logistic model. Evaluating classification and assessing the model.
4.	**Tree Based Models** ✓ Decision Trees ✓ Regression and classification Trees ✓ Random Forest	R) APM Ch. 8,& 14 (R) ISLR Ch 8 (V) from instructor	HW3 1/31	-To understand and fit tree based models both for regression and classification methods.

5.	**Neural Network** ✓ Intro. To Neural Network ✓ Fitting Neural Network Models ✓ Assessing model performance	APM Ch. 7, Ch. 13.2 and 13.5 and other supplemental materials	HW4 2/7	-To understand how to build neural network models both for regression and classifications.
6.	**Unsupervised Learning** ✓ Cluster Analysis ✓ K-means clustering ✓ Hierarchical clustering	(R) ISLR Ch 10.1,10.3, 10.5 + Other	HW5 2/14	-To understand how the observations or variables are grouped into homogeneous and distinct clusters and its applications.
7.	**Unsupervised Learning** ✓ Market Basket Analysis ✓ Basics of Association Rule ✓ Apriori algorithm	Notes and other readings from various sources	HW6 2/21 * **	To understand the frequent product sets that customers purchase together in the same "Market basket" and derive association rules from it.
			Final Exam	

*Project Presentation: **Saturday February 26th, Time 1:00pm-5:00pm**

Project Report must be uploaded to Carmen on **Monday, February 28th, at 11:59 PM.

About the Instructor

Dr. Ismael Talke is a Senior Lecturer in the Operations and Business Analytics. He comes from Miami University Ohio, department of Information Systems and Analytics, where he worked as a Visiting faculty and taught undergraduate Business Statistics, Forecasting and Analytics courses. Talke has also previously served as a faculty in Augsburg University, Minnesota and University of Asmara, Eritrea where he taught several undergraduate Statistics courses.

Talke has a PhD and MS in Statistics from Montana State University, MS in Statistics from University of KwaZulu-Natal in South Africa and BS in Mathematics from University of Asmara, Eritrea.

His professional experience includes Statistical consulting and advising undergraduate and graduate research projects. Talke is a member of the American Statistical Association.

Talke is interested in research topics in experimental design, quality control, and forecasting and data analytics. His work has been published in the Journal of Statistical Planning and Inference, International Journal of Experimental Design and Process Optimization and Academic Journal of Interdisciplinary Studies.