

**NEW JERSEY INSTITUTE OF TECHNOLOGY MARTIN TUCHMAN SCHOOL OF
MANAGEMENT**

**Data Mining and Analysis for Managers – Course Number 95136
MGMT 635-101– Fall 2023**

Instructor: Wayne G. Fox, Esq. JD/MBA,
E-mail: fox@njit.edu Telephone: 908-904-0064 cell with text
When emailing, please put “MGMT 635” in the subject line.
Duration: 5 September 2023 – 23 December 2023

Office Hours: Monday & Wednesday 11:30 am -12:30 pm and 4:30-5:30 in CAB 4027

Class: Wednesday 6:00 pm - 8:50 pm in Cullimore Lecture Hall 3

Final Exam: Week of 18 December 2023 to be scheduled.

Course Description: MGMT 635. Data Mining and Analysis. 3 credits, 3 contact hours.

This course provides an introduction to data mining with an emphasis on large scale databases as a source of knowledge generation and competitive advantage. Specific topics include: framing research questions; data modeling; inferential data mining techniques; and evaluation and deployment of data mining systems.

Through *exposure* to the theoretical base of prominent data mining methodologies along with project *experiences*, students will learn how data mining can be used in the business environment to enhance operational efficiency, Students will be equipped to explore using Teradata University Network (a free resource). Concepts covered include: Data management, conceptual model building, statistical analysis and building mining models for a variety of business applications. These objectives are accomplished through the utilization of relevant textbooks, case studies, software demonstrations and hands on data mining (individual and class projects). The student should leave this course with a sound understanding of how to utilize data mining to enhance business process efficiencies in a variety of applications. The concept of big data, machine learning and artificial intelligence will be covered throughout the semester.

Prerequisites: none (knowledge of basic computer programming is useful, not required)

Textbook: Sharda, Delen and Turban, **Business Intelligence, Analytics, and Data Science: A Managerial Perspective**, Fourth Edition, © 2008-2018 Pearson Education, ISBN 978-0-13-463328-2

Recommended supplemental books and Models: Source of some materials beyond textbook.

(1) Ajay Agrawal, Gans, & Goldfarb, **Prediction Machines, The Simple Economics of Artificial Intelligence**, Harvard Business Review Press, ©2018 (250 pages)

(2) Domingos, Pedro, **The Master Algorithm, How the Quest for the Ultimate Learning Machine will Remake Our World**, Basic Books © 2015 (330 pages, easy read)

(3) Fox, Wayne, **Natural Language Utility Models and Complex Adaptive Systems**, © 2003-2022

Required Materials: Textbook, computer, Git Hub account, NJIT supplied software and use of free online resources and training videos.

- **Canvas:** There is a Learning Management System (LMS) site (at <https://canvas.njit.edu/>) for this course on which documents will be posted as needed. Login to Canvas using your UCID and password. All announcements, assignments, changes, etc. are posted there. The student is responsible for remaining up to date at all times. The Instructor will send e-mails to the class through Canvas announcements.

Course Learning Goals and Outcomes:

- 1) Analytical and Problem Solving Skills
 - a) Students will display an ability to apply knowledge of decision science concepts to solve business problems through written analysis, use of Python based Jupyter notebooks and tests.
 - b) Students will demonstrate use of inferential and probabilistic methodologies as tools to help solve business problems in sample projects.
- 2) Communication and Information Literacy Skills
 - a) Students will be able to explain basic concepts of data mining and analysis in writing using PowerPoint slides in group video presentations.
 - b) Students will discover important sources that can be used to obtain relevant data for solving problems and how to collect data from them.
 - c) Students will participate in written class discussion questions through Canvas.
 - d) Students will work in teams to prepare project video presentation(s) on Zoom or Webex with a transcript and PowerPoint slides.
- 3) Ethical Reasoning
 - a) Students will identify ethical dilemmas that may occur in collecting personal information and determine acceptable responses to these situations.
 - b) Students will learn and discuss why interpersonal trust is more important than simply relying on written regulations and laws.
 - c) Students will observe the dangers of AI through presentation by The Center for Humane Technology <https://www.humanetech.com/>
- 4) Introduction to Dynamic Modeling using Excel, SAS, Palisade Decision Tools Suite, JMP, Minitab and other software, including Artificial Intelligence, for data mining and analysis.
 - a) Students will build integrated data analysis models using SQL, Python, R, Java or other methods as they may have experience.
- 5) Students should leave the course with an understanding of the following concepts:
 - a) -Business reporting and Analytics (BI) for decision making (Descriptive, Comparative and Predictive)
 - b) -Statistics for analytic interpretation (concepts and applied use, Excel and SAS based)
 - c) -Regression based and major Data Mining Methods (concepts and applied use)
 - d) -Process of Model Building for Problem Solving
 - e) -Solving Business verticals with Data Mining (concepts and applied use, Excel and SAS based)
- 6) IBM Skills Build Course in Artificial Intelligence with a completion certificate and badge.
- 7) Utilize ChatGPT and the new Bing with AI to explore use of scripts.

Chapter Learning Objectives reflect the goals of the textbook and are subordinate to the course goals. Be mindful of the listed objectives as you read the chapters.

Homework and Preparation: Readings from the textbook (and possibly other sources) will be assigned for each class. You should read these each week and be prepared to respond to questions about the readings. There will also be questions or problems from the textbook to answer in Canvas based quizzes for each chapter. Homework will be online and you are encouraged to upload assignments to Canvas for point credit. We will go over problems in the class as needed. If you don't understand how to do a problem, you can ask about it by email or in class or during office hours. Questions and articles may be posted on Canvas for class discussion.

Exams: There will be two or more exams, midterm (projects) and a final. Tests given in class are closed book and there may also be several shorter quizzes, which will be given in class or online using Canvas.

Project: Depending on class size, projects will be individual or group. Group size will be about 4 to 5 students. Students will be assigned Data Mining project(s) which will require utilization of multiple forms of data mining and analytic methods and an entire management/business plan. (Details to be discussed and determined with class input and interest).

Grading Policy: Final exam 270 points midterm exam (or project) 110 points, several quizzes and homework assignments (case summaries) worth 15 to 30 points each, several project assignments worth up to 250 points and group participation worth 50-100 points, class participation also earns minor points. (There are 1,000 points possible) Actual points may add to more or less than 1000 total, in which case all grades will be scaled such that the student with the highest number of points will set the 1000 point limit and all student grades adjusted ratably. All assignments are required. Late assignments will lose points, uncompleted assignments earn no points and can severely impact your grade. Canvas grades are only an approximation of your work and are not all inclusive. The instructor reserves the right to modify the weighting of the grades as to be of greatest benefit of the class as a whole. Graduate grading scale:

A	4.0	Excellent
B+	3.5	Good
B	3.0	Acceptable
C+	2.5	Marginal Performance
C	2.0	Minimum Performance
F	0.0	Failure

Academic Integrity: Students are expected to observe the NJIT Academic Integrity Code. Cheating and plagiarism will not be tolerated and may result in failing the course. ***“Academic Integrity is the cornerstone of higher education and is central to the ideals of this course and the university. Cheating is strictly prohibited and devalues the degree that you are working on. As a member of the NJIT community, it is your responsibility to protect your educational investment by knowing and following the academic code of integrity policy that is found at: <http://www5.njit.edu/policies/sites/policies/files/academic-integrity-code.pdf>.***

*Please note that it is my professional obligation and responsibility to report any academic misconduct to the Dean of Students Office. **Any student found in violation of the code by cheating, plagiarizing or using any online software inappropriately will result in disciplinary action. This may include a failing grade of F, and/or suspension or dismissal from the university.** If you have any questions about the code of Academic Integrity, please contact the Dean of Students Office at dos@njit.edu”*

Disability: Students with special needs as outlined under the Americans with Disabilities Act should first register with the Counseling Center. If you need an accommodation due to a disability please contact Scott Janz, Associate Director of the Office of Accessibility Resources and Services, Kupfrian Hall 201 to discuss your specific needs. A Letter of Accommodation Eligibility from the office authorizing student accommodations is required.

Etiquette: Students are expected to follow NJIT's Code of Student Conduct. Students should be polite, active participants and respect their instructor, classmates, as well as ideas or opinions that differ from their own. Students may not use cell phones or engage in on-line activities unrelated to class during class meetings.

Changes to this Syllabus: This syllabus may change. New versions will be announced and posted on Canvas. You are responsible for knowing what is in the latest version.

Chapter Topics will be covered approximately one chapter each week or two, in sequence:

6 SEPTEMBER 2023 CLASS

**AN OVERVIEW OF BUSINESS INTELLIGENCE, ANALYTICS, AND DATA SCIENCE
HIGHLIGHTS OF PREDICTION MACHINES AND THE MASTER ALGORITHM**

Learning Objectives for Chapter 1

- Understand today's turbulent business environment and describe how organizations survive and even excel in such an environment (solving problems and exploiting opportunities)
- Understand the need for computerized support of managerial decision making
- Recognize the evolution of such computerized support to the current state—analytics/data science
- Describe the business intelligence (BI) methodology and concepts
- Understand the various types of analytics, and see selected applications
- Understand the analytics ecosystem to identify various key players and career opportunities

13 SEPTEMBER 2023 CLASS

DESCRIPTIVE ANALYTICS I: NATURE OF DATA, STATISTICAL MODELING, AND VISUALIZATION

Learning Objectives for Chapter 2

- Understand the nature of data as it relates to business intelligence (BI) and analytics
- Learn the methods used to make real-world data analytics ready
- Describe statistical modeling and its relationship to business analytics
- Learn about descriptive and inferential statistics
- Define business reporting, and understand its historical evolution
- Understand the importance of data/information visualization
- Learn different types of visualization techniques
- Appreciate the value that visual analytics brings to business analytics
- Know the capabilities and limitations of dashboards

20 SEPTEMBER 2023 CLASS

DESCRIPTIVE ANALYTICS II: BUSINESS INTELLIGENCE AND DATA WAREHOUSING

Learning Objectives for Chapter 3

- Understand the basic definitions and concepts of data warehousing
- Understand data warehousing architectures
- Describe the processes used in developing and managing data warehouses
- Explain data warehousing operations
- Explain the role of data warehouses in decision support
- Explain data integration and the extraction, transformation, and load (ETL) processes
- Understand the essence of business performance management (BPM)
- Learn balanced scorecard and Six Sigma as performance measurement systems

27 SEPTEMBER CLASS

CLASS PROJECT 1

4 OCTOBER 2023

PREDICTIVE ANALYTICS I: DATA MINING PROCESS, METHODS, AND ALGORITHMS

Learning Objectives for Chapter 4

- Define data mining as an enabling technology for business analytics
- Understand the objectives and benefits of data mining
- Become familiar with the wide range of applications of data mining
- Learn the standardized data mining processes
- Learn different methods and algorithms of data mining
- Build awareness of the existing data mining software tools
- Understand the privacy issues, pitfalls, and myths of data mining

11 OCTOBER 2023 CLASS

MIDTERM EXAM 1

18 OCTOBER 2023 CLASS

PREDICTIVE ANALYTICS II: TEXT, WEB, AND SOCIAL MEDIA ANALYTICS

Learning Objectives for Chapter 5

- Describe text analytics and understand the need for text mining
- Differentiate among text analytics, text mining, and data mining
- Understand the different application areas for text mining
- Know the process of carrying out a text mining project
- Appreciate the different methods to introduce structure to text-based data
- Describe sentiment analysis
- Develop familiarity with popular applications of sentiment analysis
- Learn the common methods for sentiment analysis
- Become familiar with speech analytics as it relates to sentiment analysis

25 OCTOBER 2023 CLASS

CLASS PROJECT 2

1 NOVEMBER 2023

CLASS PROJECT 2 PRESENTATIONS AND STUDENT GRADING

8 NOVEMBER 2023

PRESCRIPTIVE ANALYTICS: OPTIMIZATION AND SIMULATION

Learning Objectives for Chapter 6

- Understand the applications of prescriptive analytics techniques in combination with reporting and predictive analytics
- Understand the basic concepts of analytical decision modeling
- Understand the concepts of analytical models for selected decision problems, including linear programming and simulation models for decision support
- Describe how spreadsheets can be used for analytical modeling and solutions
- Explain the basic concepts of optimization and when to use them
- Describe how to structure a linear programming model
- Explain what is meant by sensitivity analysis, what-if analysis, and goal seeking
- Understand the concepts and applications of different types of simulation
- Understand potential applications of discrete event simulation

15 NOVEMBER 2023 CLASS

BIG DATA CONCEPTS AND TOOLS

Learning Objectives for Chapter 7

- Learn what Big Data is and how it is changing the world of analytics
- Understand the motivation for and business drivers of Big Data analytics
- Become familiar with the wide range of enabling technologies for Big Data analytics
- Learn about Hadoop, MapReduce, and NoSQL as they relate to Big Data analytics
- Compare and contrast the complementary uses of data warehousing and Big Data
- Become familiar with the vendors of Big Data tools and services
- Understand the need for and appreciate the capabilities of stream analytics
- Learn about the applications of stream analytics

29 NOVEMBER 2023 CLASS

FUTURE TRENDS, PRIVACY AND MANAGERIAL CONSIDERATIONS IN ANALYTICS

Learning Objectives for Chapter 8

- Explore some of the emerging technologies that may impact analytics, business intelligence (BI), and decision support
- Describe the emerging Internet of Things (IoT) phenomenon, potential applications, and the IoT ecosystem
- Describe the current and future use of cloud computing in business analytics
- Describe how geospatial and location-based analytics are assisting organizations
- Describe the organizational impacts of analytics applications
- List and describe the major ethical and legal issues of analytics implementation
- Identify key characteristics of a successful data science professional

6 DECEMBER 2023

MIDTERM EXAM 2 / PROJECT 3

13 DECEMBER 2023 LAST REGULAR CLASS – GROUP PRESENTATIONS STUDENT GRADED

14/15 NOVEMBER 2023 READING DAY

20 NOVEMBER 2023 FINAL EXAM

The Doctoral Level parallel course covers similar topics in greater depth:

MGMT 735. Deep Learning in Business. 3 credits, 3 contact hours.

Prerequisites: [FIN 620](#) or instructor's approval or advanced graduate standing. This course provides an in-depth study of data mining and machine learning, with a focus on business applications. As the business market becomes increasingly complicated and depends on data, analysts and fund managers must make better and faster decisions using available data. Data mining and machine learning make use of powerful tools and techniques to unlock the value inherent in available market data and routinely help managers uncover hidden patterns and correlations in data and gain insights to improve the decision-making in the market. The course is practice-oriented and develops the required skills to apply machine learning in the stock market and other business areas. Students will better understand the techniques for data mining and machine learning as well as gain hands-on knowledge of the contemporary analysis tools of data mining and machine learning. The course will enable students to better understand the major concepts, approaches, and techniques for data mining and machine learning. The included learning material provides adequate technical depth for students to know how data-driven technologies work. Coverage includes data mining and machine learning processes, methods, and techniques; the role and management of data; tools and metrics; and integration with Big Data.

MGMT 620. Management of Technology. 3 credits, 3 contact hours.

This course explores technology as a main component of an organizational entity. The generation, development, and implementation of technology are outlined. The influence of technology on global competitiveness is also discussed.

AWS Certified Cloud Practitioner Training 2020 - Full Course

<https://youtu.be/3hLmDS179YE> 3 hours and 58 minutes (to be discussed)

[Teradata University - This will allow you to sign into the application with your username or userID followed by your application password](#)

Additional Resources Used in Preparation for this Class:

- 1) **Real World Data Mining, Applied Business Analytics and Decision Making**, by Dursun Delen © 2014 Pearson Education. (do not purchase additional texts, use library) (Amazon \$50 11/30/21)
- 2) **Business Intelligence, Analytics and Data Science: A Managerial Perspective**, by Ramesh Sharda, Dursun Delen and Efraim Turban 4th ed © 2017 Pearson Education (Amazon \$112 new to \$65 used)
- 3) **Analytics, Data Science, & Artificial Intelligence Systems for Decision Support** 11th Edition by Ramesh Sharda; Dursun Delen; Efraim Turban and Publisher Pearson. Copyright © 2020
- 4) **Getting Started with Artificial Intelligence: A Practical Guide to Building Enterprise Applications** by Tom Markiewicz and Josh Zheng Copyright © 2018 International Business Machines Corporation. Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472. (Free PDF)
- 5) **Prediction Machines The Simple Economics of Artificial Intelligence** by Ajay Agrawal; Joshua Gans; Avi Goldfarb, Publisher: © 2018 Harvard Business Review Press (Amazon \$16 - \$26.50)
- 6) **Power and Prediction The Disruptive Economics of Artificial Intelligence** by Ajay Agrawal; Joshua Gans; Avi Goldfarb, Publisher: © 2022 Harvard Business Review Press Amazon \$27)
- 7) **Data Analytics, Data Visualization & Communicating Data** (3 books in one) by Elizabeth Clarke © 2022 (Amazon \$30.46)
- 8) **The Master Algorithm, How the Quest for the Ultimate Learning Machine will Remake Our World** by Pedro Domingos © 2015 Basic Books (Amazon \$10.94)
- 9) **Python for Programmers with Introductory AI Case Studies**, by Paul Deitel & Harvy Deitel © 2019 Pierson Education (Amazon \$60)
- 10) **The Age of AI and our Human Future**, by Henry A. Kissenger, Eric Schmidt and Daniel Huttenlocher © 2021 Back Bay Books (Amazon \$17)
- 11) **Predictive Analytics – Data Mining, Machine Learning and Data Science for Practioners**, 2nd ed by Dursun Delen © 2021 Person Education (Amazon \$46.07)
- 12) **Age of Invisible Machines, A Practical Guide to Creating a Hyperautomated Ecosystem of Intelligent Digital Workers**, by Robb Wilson with Josh Tyson © 2023 Wiley (Amazon \$21.97)
- 13) **Machine Learning for Algorithmic Trading**, by Stefan Jansen © 2020 Packt Publishing (Amazon \$45)
- 14) **The Quantum Brain**, by Amit Goswami, PhD and Vlaentina R. Onisor, MD © 2021Bluerose Publishers (Amzn \$14)
- 15) **Mental Models – Toward a Cognitive Science of Language, Inference and Consiousness**, by P. N. Johnson Laird © 1983 Harvard University Press

- 16) **Introduction to Machine Learning**, by Michael L. Littman, PhD, Brown University, The Great Courses video and Git Hub Jupyter Notebooks
- 17) **How You Decide: The Science of Human Decisionmaking**, by Prof Ryan Hamilton, The Great Courses video
- 18) **Big Data: How Data Analytics is Transforming the World**, by Tim Chartier PhD, The Great Courses video
- 19) **Language and the Mind**, by Spenser Kelley PhD, The Great Courses video
- 20) **Mathematical Decisionmaking: Predictive Models and Optimization**, Scott P. Stevens PhD, James Madison University, The Great Courses video

Course Reference Materials

		Title						Pedros Domingos	Mitchell	Russell & Norvig	Charniak
		INTRODUCTION TO MACHINE LEARNING						The Master Algorithm	Machine Learning	Artificial Intelligence	Intro to Deep Learning
		Michael L Littman, PhD, Brown University									
Lesson		LESSON GUIDES		Colab	Colab Aux		Chapters	Chapter	Sections	Chapters	
1	Lesson 01	Telling the Computer What We Want 4	4	14			1-2	1	19.1-19.2		
2	Lesson 02	Starting with Python Notebooks and Colab 16	16	35	L02.ipynb						
3	Lesson 03	Decision Trees for Logical Rules 24	24	46	L03.ipynb	L03aux.ipynb	3	3	19.3	1-2	
4	Lesson 04	Neural Networks for Perceptual Rules 27	27	63	L04.ipynb		4	4	21.1		
5	Lesson 05	Opening the Black Box of a Neural Network 32	32	78	L05.ipynb				19.6		
6	Lesson 06	Bayesian Models for Probability Prediction 36	36	96	L06.ipynb		6	6	Ch. 12, 20.2		
7	Lesson 07	Genetic Algorithms for Evolved Rules 42	42	113	L07.ipynb	L07aux.ipynb	5	9			
8	Lesson 08	Nearest Neighbors for Using Similarity 47	47	131	L08.ipynb	L08aux.ipynb	7	8	19.7		
9	Lesson 09	The Fundamental Pitfall of Overfitting 52	52	148	L09.ipynb	L09aux.ipynb		7	19.5		
10	Lesson 10	Pitfalls in Applying Machine Learning 57	57	166	L10.ipynb				19.9		
11	Lesson 11	Clustering and Semi-Supervised Learning 64	64	187	L11.ipynb		8		21.7	7	
12	Lesson 12	Recommendations with Three Types of Learning 72	72	205	L12.ipynb				Ch. 17, 19.8		
13	Lesson 13	Games with Reinforcement Learning 79	79	224	L13.ipynb			13	Ch 5 & 22, 21.8.3		
14	Lesson 14	Deep Learning for Computer Vision 86	86	242	L14.ipynb	L14aux.ipynb			21.3 & 21.8.1	3	
15	Lesson 15	Getting a Deep Learner Back on Track 91	91	258	L15.ipynb				19.4		
16	Lesson 16	Text Categorization with Words as Vectors 97	97	277	L16.ipynb				21.8.2, 24.1	4	
17	Lesson 17	Deep Networks That Output Language 105	105	297	L17.ipynb				Ch. 23-24	5	
18	Lesson 18	Making Stylistic Images with Deep Networks 112	112	318	L18.ipynb	L18aux.ipynb					
19	Lesson 19	Making Photorealistic Images with GANs 122	122	337	L19.ipynb	L19aux.ipynb					
20	Lesson 20	Deep Learning for Speech Recognition 128	128	357	L20.ipynb						
21	Lesson 21	Inverse Reinforcement Learning from People 136	136	375	L21.ipynb				Ch. 27, 21.8.3, 22.6	6	
22	Lesson 22	Causal Inference Comes to Machine Learning 143	143	395	L22.ipynb						
23	Lesson 23	The Unexpected Power of Over-Parameterization 150	150	415	L23.ipynb						
24	Lesson 24	Protecting Privacy within Machine Learning 159	159	435	L24.ipynb						
25	Lesson 25	Mastering the Machine Learning Process 167	167	457	L25.ipynb				Ch. 6		

HOW THE QUEST FOR THE ULTIMATE LEARNING MACHINE WILL REMAKE OUR WORLD - PEDROS DOMINGOS

		Pedros Domingos						Pedros Domingos	Mitchell	Russell & Norvig	Charniak
		INTRODUCTION TO MACHINE LEARNING						The Master Algorithm	Machine Learning	Artificial Intelligence	Intro to Deep Learning
		Michael L Littman, PhD, Brown University									
How the Quest for the Ultimate Learning Machine will Remake Our World	Lesson	LESSON GUIDES		Colab	Colab Aux		Chapters	Chapter	Sections	Chapters	
Basic Books © 2015											
Prologue											
Chapter 1: The Machine-Learning Revolution	1	22	1	Lesson 01 Telling the Computer What We Want 4	4	14		1-2	1	19.1-19.2	
Chapter 2: The Master Algorithm	23	34		Lesson 01 Telling the Computer What We Want 4	4	14		1-2	1	19.1-19.2	
Chapter 3: Hume's Problem of Induction	57	36	3	Lesson 03 Decision Trees for Logical Rules 24	24	46	L03.ipynb	L03aux.ipynb	3	3	19.3
Chapter 4: How Does Your Brain Learn?	93	28	4	Lesson 04 Neural Networks for Perceptual Rules 27	27	63	L04.ipynb		4	4	21.1
Chapter 5: Evolution: Nature's Learning Algorithm	121	22	7	Lesson 07 Genetic Algorithms for Evolved Rules 42	42	113	L07.ipynb	L07aux.ipynb	5	9	
Chapter 6: In the Church of the Reverend Bayes	143	34	6	Lesson 06 Bayesian Models for Probability Prediction 36	36	96	L06.ipynb		6	6	Ch. 12, 20.2
Chapter 7: You Are What You Resemble	177	26	8	Lesson 08 Nearest Neighbors for Using Similarity 47	47	131	L08.ipynb	L08aux.ipynb	7	8	19.7
Chapter 8: Learning Without a Teacher	203	32	11	Lesson 11 Clustering and Semi-Supervised Learning 64	64	187	L11.ipynb		8		21.7
Chapter 9: The Pieces of the Puzzle Fall into Place	235	28									
Chapter 10: This is the World on Machine Learning	263	28									
Epilogue	291	4									
Acknowledgments	295	2									
Further Readings	297	18									
Index	315	14									
End of Volume	329										