

DATA MINING FOR BUSINESS INTELLIGENCE

MIS 373 17 (Unique 03705), MKT 372 (Unique 04935)

T Th 2:00-3:30 p.m. (UTC 1.130)

Instructor: Professor Anitesh Barua

Office: CBA 5.232

Email: Anitesh.Barua@mcombs.utexas.edu

Office hours: T Th 9:00 – 10:30 a.m. (or by appointment)

Course Overview

The rapid proliferation of the Internet and related technologies has created an unprecedented opportunity for enterprises to collect massive amounts of data regarding customers and all aspects of their business operations. Yet the reality is that most organizations today are (i) “data rich” but “information and knowledge poor”, and (ii) not harnessing the full potential of their data, which is perhaps the second most important asset after human capital. Internet based applications such as social media, website usage tracking and online reviews as well as more traditional technology applications like RFID, Supply Chain Management (SCM), Enterprise Resource Planning (ERP) and Customer Relationship Management (CRM) provide access to vast amounts of data regarding customers, suppliers, competitors as well as a firm’s own activities and business processes. Being able to unlock the insights and knowledge trapped in such raw data constitutes a key lever for competitive advantage in hypercompetitive business environments.

This course is designed to showcase the virtually unlimited opportunities that exist today to leapfrog the competition by leveraging the data that organizations routinely collect every day, but which they hardly use strategically to make decisions at various points in the value chain. Students will be exposed to a wide gamut of issues related to data analytics and business intelligence, including the strategic aspects of big and better data as well as the details of analytical methods and data mining and visualization tools such as XLMiner and Node XL.

Learning Objectives

This course is especially valuable to students contemplating careers in business analytics, marketing, prediction modeling, consulting and general management. Students taking this course will develop expertise in the following areas:

1. Strategic aspects and business value of data analytics
2. Data capture, validation, reduction, analysis, insights and recommendations
3. Practical analytical and technical skills that differentiates you in any modern enterprise
4. In depth expertise in techniques and methods of classification, prediction, and association
5. Real world data analytic and business intelligence applications

Students are not required to have a deep knowledge of statistics (though a basic understanding is necessary) or technical ability in programming languages and software applications. The content of this course is presented in an intuitive format with emphasis on the connection between data and business strategies. A key feature of this course is the use of XLMiner (an Excel add-in) for data mining and NodeXL (an open-source template for Microsoft Excel 2007 and 2010) for the analysis of social media networks.

Course Material

Textbooks

1. "Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner"
by Galit Shmueli, Nitin R. Patel, Peter C. Bruce
Publisher: Wiley; 2 edition (October 26, 2010)
ISBN-10: 0470526823
ISBN-13: 978-0470526828
2. "Analyzing Social Media Networks with NodeXL: Insights from a Connected World"
by Derek Hansen, Ben Shneiderman and Marc A. Smith
Publisher: Morgan Kaufmann; 1 edition (September 10, 2010)
ISBN-10: 0123822297
ISBN-13: 978-0123822291

Readings list (course packet available at IT Copy & Printing, 512 W Martin Luther King Jr, Austin, TX 78701, Tel: 512-476-6662)

Articles

1. “Big Data, Analytics and the Path From Insights to Value,” by Steve LaValle, Eric Lesser, Rebecca Shockley, Michael S. Hopkins, Nina Kruschwitz HBS SMR372-PDF-ENG.
2. “Architecture of Business Intelligence: Aligning a Robust Technical Environment with Business Strategies,” by Thomas H. Davenport, Jeanne G. Harris, HBS 2202BC-PDF-ENG
3. “A Step-By-Step Guide to Smart Business Experiments,” by Eric T. Anderson, Duncan Simester, HBS R1103H-PDF-ENG
4. “Embed Analytics in Business Processes: A How-To Guide,” by Thomas H. Davenport, Jeanne G. Harris, , HBS 5751BC-PDF-ENG

Cases

1. “Business Intelligence Software at SYSCO,” by Andrew McAfee, Alison Berkley Wagonfeld, HBS 604080-PDF-ENG
2. “Harrah's Entertainment Inc.: Real-Time CRM in a Service Supply Chain,” by Hau Lee, Seungjin Whang, Kamram Ahsan, Earl Gordon, Amir Faragalla, Asha Jain, Abid Mohsin, Shi Guangyu, Guangyu Shi, HBS GS50-PDF-ENG
3. “Harrah's Entertainment, Inc.” by Rajiv Lal, Patricia Martone Carrolo, HBS 502011-PDF-ENG
4. “Netflix Leading with Data: The Emergence of Data-Driven Video,” by Russell Walker, Mark Jeffery, Linus So, Sripad Sriram, Jon Nathanson, Joao Ferreira, HBS KEL473-PDF-ENG
5. “Testing, Monitoring, and Adjusting Strategic Objectives Through Data Analytics at Northwestern Mutual,” by Anne Field, HBS B1107B-PDF-ENG

Grading

Your course grade will be based on the following:

Item	Date due	Weight
Individual assignments	2/9, 3/1, 3/27, 4/10, 4/26	35%
Group project	Final presentation on 5/1 and 5/3	20%
Midterm	3/6	15%
Take home final	Handed out May 4, due May 11 by 11:59 p.m.	20%
Class participation		10%

Class participation: 10%

In this class much of the learning is dependent on the accessing the combined knowledge and experience of the group. It is everyone's job to keep the discussion productive and moving forward. In evaluating your class participation grade, I take the following into consideration:

- useful arguments expressed coherently and succinctly
- good analysis supported by case facts or your own experience
- relevance to previous contributions, i.e. ability to listen and build on what others say
- constructive disagreement
- regard, respect and acknowledgment of others' contributions
- readiness to contribute to class discussions

Individual Assignments (35%)

There will be 5 individual assignments during the course. The schedule of assignments is as follows:

1. Harrah's (case study, readings packet): 2/9
2. Charles Book Club (p. 367 of Data Mining textbook): 3/1
3. German Credit (p. 375 of Data Mining textbook): 3/27
4. "Segmenting Consumers of Bath Soap" (p. 383): 4/10
5. "Cosmetics Purchases" (p. 277): 4/26

Group project (20%)

Students will work in groups of **five** on a semester-long data mining and business intelligence project dealing with real world data. You will be responsible for forming your own team. Topics can vary widely depending on student experience and interest, and can include areas such as healthcare (e.g., factors that drive operating efficiency and quality of care), finance and financial services (e.g., trading strategies, predicting loan defaults), and electronic commerce (e.g., online customer acquisition/retention, customization and pricing strategies). Groups are responsible for initiating contact with organizations or sources of data. Groups will make three presentations during the semester:

2/21: Groups will present their proposed topics and initial progress.

3/22: Groups will present the status of their projects.

5/1 and 5/3: All student groups should be ready to present on 5/1. Half the groups will be chosen randomly to present their studies. The remaining groups will present on 5/3.

PowerPoint slides + details of analysis will be submitted by all groups by the beginning of class on 5/1.

Course Agenda

Date	Topic	Readings
1/17	Data mining, business strategies and value	<ol style="list-style-type: none"> 1. <u>"Big Data, Analytics and the Path From Insights to Value,"</u> by <i>Steve LaValle, Eric Lesser, Rebecca Shockley, Michael S. Hopkins, Nina Kruschwitz</i> 2. <u>"Architecture of Business Intelligence: Aligning a Robust Technical Environment with Business Strategies,"</u> by <i>Thomas H. Davenport, Jeanne G. Harris</i>
1/19	Transforming business processes and operations with data analytics	<ol style="list-style-type: none"> 1. <u>"A Step-By-Step Guide to Smart Business Experiments,"</u> by <i>Eric T. Anderson, Duncan Simester, HBS R1103H-PDF-ENG</i> 2. <u>"Testing, Monitoring, and Adjusting Strategic Objectives Through Data Analytics at Northwestern Mutual,"</u> by <i>Anne Field, HBS B1107B-PDF-ENG</i> 3. <u>"Embed Analytics in Business Processes: A How-To Guide,"</u> by <i>Thomas H. Davenport, Jeanne G. Harris, HBS 5751BC-PDF-ENG</i>
1/24	Best practices in data analytics and business intelligence	<ol style="list-style-type: none"> 1. <u>"Netflix Leading with Data: The Emergence of Data-Driven Video,"</u> by <i>Russell Walker, Mark Jeffery, Linus So, Sripad Sriram, Jon Nathanson, Joao Ferreira, HBS KEL473-PDF-ENG</i> 2. <u>"Business Intelligence Software at SYSCO,"</u> by <i>Andrew McAfee, Alison Berkley Wagonfeld, HBS 604080-PDF-ENG</i>
1/26	The data mining process	<ol style="list-style-type: none"> 1. Data Mining for Business Intelligence: Chapters 1, 2 2. Install XLMiner on your computer
1/31	The data mining process	Data Mining for Business Intelligence: Chapter 2
2/2	Dimension reduction	Data Mining for Business Intelligence: Chapter 4

2/7	Performance evaluation	Data Mining for Business Intelligence: Chapter 5
2/9	Best practices in data analytics: The case of Harrah's	<p>1. Harrahs "<u>Harrah's Entertainment Inc.: Real-Time CRM in a Service Supply Chain</u>," by <i>Hau Lee, Seungjin Whang, Kamram Ahsan, Earl Gordon, Amir Faragalla, Asha Jain, Abid Mohsin, Shi Guangyu, Guangyu Shi</i>, HBS GS50-PDF-ENG</p> <p>2. "<u>Harrah's Entertainment, Inc.</u>" by <i>Rajiv Lal, Patricia Martone Carolo</i>, HBS 502011-PDF-ENG</p> <p>Assignment #1 due by the beginning of class: Harrah's case studies</p>
2/14	Classification	Data Mining for Business Intelligence: Chapters 7, 8
2/16	Classification	Data Mining for Business Intelligence: Chapter 9
2/21	Group Project Proposals	Student groups will present their topic and initial progress
2/23	Classification	Data Mining for Business Intelligence: Chapter 10
2/28	Neural Networks	Data Mining for Business Intelligence: Chapter 11
3/1	Discriminant analysis	Data Mining for Business Intelligence: Chapter 12
		Assignment #2 due: "Charles Book Club"
3/6	Midterm	
3/8	Association rules	Data Mining for Business Intelligence: Chapter 13
3/20	Guest speaker	TBA
3/22	Project Review	Student groups will present their project status
3/27	Association rules	Data Mining for Business Intelligence: Chapter 13
		Assignment #3 due: "German Credit"
3/29	Clustering	Data Mining for Business Intelligence: Chapter 14
4/3	Visualization of data	Data Mining for Business Intelligence: Chapter 3
4/5	Visualization of data	Data Mining for Business Intelligence: Chapter 3
4/10	Extracting business intelligence from social media	Analyzing Social Media Networks with NodeXL: Ch. 1, 3
		Assignment #4 due: "Segmenting Consumers of Bath Soap"

4/12	Tools for analyzing social media networks	Analyzing Social Media Networks with NodeXL: Chapters 4, 5, 6, 7
4/17	Guest speaker	TBA
4/19	Analyzing Twitter	Analyzing Social Media Networks with NodeXL, Chapter 10
4/24	Analyzing Facebook	Analyzing Social Media Networks with NodeXL, Chapter 11
4/26	Review and course summary	Assignment #5 due: "Cosmetics Purchases" 13.3, p. 277 of Data Mining book.
5/1	Project presentations	All student groups should be ready to present on 5/1. Half the groups will be chosen randomly to present their studies. PowerPoint slides will be submitted by all groups by the beginning of class.
5/3	Project presentations	Remaining half of the groups will present their studies.
5/4	Take home final handed out	
5/11	Take home final due by 11:59 p.m.	