

Syllabus

Course Number: MSDS 682

Course Title: Text Analytics

Course Description:

Investigates linguistic, statistical, and machine learning techniques for modeling the information in textual sources. Includes information retrieval, natural language processing, text classification, and sentiment analysis and the software systems for performing these analyses.

Prerequisite Courses:

MSDS 650 Data Analytics

Course Overview

In this course, you will learn the fundamental concepts of text analytics and perform text analytics on different applications.

Key concepts to be covered in this course include:

- Text analytics concepts and applications
- Fundamental of Information retrieval and natural language processing
- Text analytics framework
- Theoretical techniques and applications in text analytics (e.g. social media)
- R packages and commands to perform text analytics

Course Outcomes:

Upon completion of this course, learners should be able to:

- Explain text analytics framework.
- Analyze on your own text data.
- Measure model performance with appropriate metrics measurements.
- Interpret the results, gain insight and recommend possible actions.

Course Materials:

Manning, C. D., Raghavan, P., and Schütze, H. 2008. *Introduction to Information Retrieval*. Cambridge University Press, 2008. Retrieved from: <http://www-nlp.stanford.edu/IR-book/>.

Miner, G., Delen, D., Elder, J., Fast, A., Hill, T., and Nisbet, A. R. (2012). *Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications*. Elsevier Inc. Available online <http://www.gbv.de/dms/ilmenau/toc/668584769.PDF>.

Required Texts:

None

Required Resources:

None

Technology Tools:

- The comprehensive R Archive Network (2016). Download and install R. (Version 3.3.2) [Software] Available from: <https://cran.r-project.org>.
- R Studio (2016). Download and install R Studio. (Version 1.0.136) [Software]. Available from: <https://www.rstudio.com/products/rstudio/download/>.
- Data Camp: Text Mining: Bag of Words (video). Retrieved from <https://www.datacamp.com/courses/intro-to-text-mining-bag-of-words>.

Optional Materials:

Munzert, S., Rubba, C., Meißner, P., and Nyhuist, D. (2015). *Automated Data Collection with R: A practical guide to web scraping and text mining*. John Wiley & Sons. (E-book). Retrieved from <http://kek.ksu.ru/eos/WM/AutDataCollectR.pdf>.

The following e-books are available to you for free through the Regis Library collections. Go to <http://libguides.regis.edu/az.php?a=s> and select Safari Books Online. You will be required to log in using your Regis credentials. **NOTE: Limited to 3 concurrent users.** Please logout when you are through.

Kumar, A., and Paul, A. (2016). *Mastering Text Mining with R*. Packt Publishing. (E-book).

Ravindran, S. K., and Garg, V. (2015). *Mastering Social Media Mining with R*. Packt Publishing (E-book).

Pre-Assignment:

Online Format: Sign on to WorldClass (LMS) and become familiar with the course navigation. See Course Assignments and Activities table below.

Classroom-based Format: Instructor will make assignments.

Pre-Assignment Due Dates:

Classroom-based Format: This assignment is due the first night of class.

Online Format: The instructor will specify the due date for this assignment.

Course Assignments and Activities:

Week	Topic	Readings	Activities Assignments and Percentages
1	Text Analytics	From the Expert Reading List for week 1	Introductions Discussion Questions/threads (1.5%) Exercise (8%)
2	Informational Retrieval	From the Expert Reading List for week 2	Discussion Questions/threads (1.5%) Exercise (8%)
3	Natural Language Processing (NLP)	From the Expert Reading List for week 3	Discussion Questions/threads (1.5%) Project (12%)
4	Text Categorization	From the Expert Reading List for week 4	Discussion Questions/threads (1.5%) Project (12%)
5	Text Clustering	From the Expert Reading List for week 5	Discussion Questions/threads (1.5%) Project (12%)
6	Social media	From the Expert Reading List for week 6	Discussion Questions/threads (1.5%) Project (12%)
7	Sentiment Analysis	From the Expert Reading List for week 7	Discussion Questions/threads (1.5%) Project (12%)
8	Other applications – health care	From the Expert Reading List for week 8	Discussion Questions/threads (1.5%) Project (12%)
	Total		100%

Student Evaluation Grid:

Assignments	Weighted Percentage
Discussion Questions (8 at 1.5% each)	12%
Exercises (2 exercises at 8% each)	16%
Projects (6 projects at 12% each)	72%
TOTAL	100 %

CC&IS Grading Scale

Letter Grade	Percentage	Grade Point
A	93 to 100	4.00
A-	90 to less than 93	3.67
B+	87 to less than 90	3.33
B	83 to less than 87	3.00
B-	80 to less than 83	2.67
C+	77 to less than 80	2.33
C	73 to less than 77	2.00
C-	70 to less than 73	1.67
D+	67 to less than 70	1.33
D	63 to less than 67	1.00
D-	60 to less than 63	.67
F	Less than 60	0

Additional information about grading can be found in the latest edition of the University Catalog, available at <http://www.regis.edu/Academics/Course%20Catalog.aspx>.

CC&IS Policies and Procedures

Each of the following CC&IS Policies & Procedures is incorporated here by reference. Students are expected to review this information each term, and agree to the policies and procedures as identified here and specified in the latest edition of the University Catalog, available at <http://www.regis.edu/Academics/Course%20Catalog.aspx> or at the link provided.

- The CC&IS Academic Integrity Policy.
- The Student Honor Code and Student Standards of Conduct.
- Incomplete Grade Policy, Pass / No Pass Grades, Grade Reports.
- The Information Privacy policy and FERPA. For more information regarding FERPA, visit the [U.S. Department of Education](#).
- The HIPAA policies for protected health information. The complete Regis University HIPAA Privacy & Security policy can be found here: <http://www.regis.edu/About-Regis-University/University-Offices-and-Services/Auxiliary-Business/HIPAA.aspx>.
- The Human Subjects Institutional Review Board (IRB) procedures. More information about the IRB and its processes can be found here: <http://regis.edu/Academics/Academic-Grants/Proposals/Regis-Information/IRB.aspx>.

The CC&IS Policies & Procedures Syllabus Addendum summarizes additional important policies including, Diversity, Equal Access, Disability Services, and Attendance & Participation that apply to every course offered by the College of Computer & Information Sciences at Regis University. A copy of the CC&IS Policies & Procedures Syllabus Addendum can be found here: <https://in2.regis.edu/sites/ccis/policies/Repository/CCIS%20Syllabus%20Addendum.docx>.