Web Analytics and Mining
MET CS 688 C1

Instructor:
Zlatko Vasilkoski
email: zlatko@bu.edu
Office hours: by appointment

Class Time: Wednesdays from 6:00pm to 8:45pm,
Location: SHA 111
Semester Starts: 1/24/2018 to 4/30/2018.

Course Description

The Web Analytics and Mining course covers the areas of web analytics, text mining, web mining, and practical application domains. The web analytics part of the course studies the metrics of web sites, their content, user behavior, and reporting. Google analytics tool is used for collection of web site data and doing the analysis. The text mining module covers the analysis of text including content extraction, string matching, clustering, classification, and recommendation systems. The web mining module studies how web crawlers process and index the content of web sites, how search works, and how results are ranked. In addition, certain aspects of cloud computing will be illustrated, and you will create probably your first Internet of Things (IoT) analytics. Application areas mining the social web and game metrics will be extensively investigated.

Please note that this is a laboratory course in which in–class use of computers is required. The class is not scheduled in a lab with computers, thus laptops are required. Also note that substantial programming experience with R is needed for this course. Some of the assignments involve creation of Shiny web apps that you may run on mobile devices and data analytics involving optimization of neural networks.

Required Course Prerequisites

CS544 (Foundations of Analytics) or CS555 (Data Analysis and Visualization).

Course Grading Policy

The course grade will be based on active class participation and quizzes/lab projects (10%), assignments (20%), midterm exam (35%), and a term project (35%). Assignments are expected to be submitted by their respective due dates. Late submissions carry a penalty.

Course Topics

Module 1 - Web Analytics
- Metrics
- Key performance indicators
- Referrers and visitors
- Identifying important pages
- Web site visibility

Module 2 - Web Analytics Tools
- Using Google Analytics
- Collecting data with Google Analytics
- Dimensions and Segmentation
- Flow visualization, navigating reports

Module 3 - Text Mining
- Preprocessing and content extraction
- Searching and fuzzy string matching
- Clustering text
- Classification, categorization, and tagging
- Question answering systems

Module 4 - Web Mining
- Web Crawlers, Indexing
- Searching, precision and recall
- Ranking
- IoT (Internet of Things)

Module 5 - Applications - Mining the Social Web
- Twitter - trending topics, Facebook - Social Graph API
- LinkedIn - Clustering colleagues, Google - Document Similarity

Module 6 - Applications - Game Analytics

- Game metrics, telemetry, and analytics
- Telemetry collection and tools
- Game data analysis and visualization
- Case studies

Reference Textbooks

Michael Beasley, Practical Web Analytics for User Experience: How Analytics can help you Understand your Users, Morgan Kaufmann, 2013.

Grant S. Ingersoll, Thomas S. Morton, and Andrew L. Farris, Taming Text: How to find, organize, and manipulate It, Manning Publications, 2013.

Matthew A. Russell, Mining the Social Web, 2nd Edition, O'Reilly, 2013

Magy Seif El-Nasr, Anders Drachen, Alessandro Canossa, eds., Game Analytics: Maximizing the Value of Player Data, Springer, 2013.

Barney L. Capehart, Ph.D., C.E.M., Michael R. Brambley, Ph.D., Automated Diagnostics and Analytics for Buildings, Fairmont Press, 2014.

Course Web Site

http://learn.bu.edu

Student Conduct Code

http://www.bu.edu/met/for-students/met-policies-procedures-resources/academic-conduct-code/