

**Boston University** Metropolitan College

COMPUTER SCIENCE DEPARTMENT

**Web Analytics and Mining**

**MET CS 688**

**Instructor:**

Zlatko Vasilkoski

email: zlatko@bu.edu

Office hours: by appointment

**Class Time:** Wednesday from 6:00pm to 9:00pm, FLR 265 (808 Comm Ave. PC labs)

Semester Starts: Wednesday, September 2. The last day of classes is Thursday, December 10, 2015. Final exams date will be assigned between December 15th and December 19th.

**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. Application areas mining the social web and game metrics will be extensively investigated. Laboratory Course.

**Course Prerequisites**

MET CS 544 - Foundations of Analytics

**Course Grading Policy**

The course grade will be based on active class participation and quizzes (10%), assignments (30%), midterm exam (30%), and a term project (30%). 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

**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.

**Course Web Site**

[**http://learn.bu.edu**](http://learn.bu.edu)

**Student Conduct Code**

[**http://www.bu.edu/met/for-students/met-policies-procedures-resources/academic-conduct-code/**](http://www.bu.edu/met/for-students/met-policies-procedures-resources/academic-conduct-code/)