Data Analytics in Social Media
Fall 2016

Instructor Basic Information
Name: Liang Zhao
Office: Room 5343, Volgeneau Engineering Building
Phone: TBD
Email: lzhao9@gmu.edu
Office Hours: Thursday 3:00pm-5:00pm

Course Basic Information
Course #: AIT 690
Course Title: AIT 690 – Data Analytics in Social Media
Term: Fall 2016
Time: Friday, 16:30-19:10
Building: Robinson Hall
Room: A105

Course Catalog
Social media has experienced a fast growth during the past decade. Millions of users in the sites such as Twitter, Facebook, and YouTube have been generating and sharing variety of content including texts, images, videos, and other metadata. In addition, social media can be treated as social sensor to reflect different aspects of the society. Data analytics in social media have enormous significance for many applications like public sentiment, business intelligence, and disaster management. Social media data has important characteristics including dynamics, heterogeneity, noisiness, timeliness, big volume, and network properties. These characteristics cause various new challenges and hence invoke many interesting research topics which will be the focus of this course.

Learning Objectives
This course aims to introduce the necessary theories and the state-of-the art techniques in Web mining, networks analysis, information retrieval, and predictive modeling to study emerging problems with social media. These problems include information diffusion, recommendations, behavior analysis, and event analytics in social media. The ultimate goal of this course is to sharpen problem solving skills of the graduate students, and prepare them with this unique set of expertise for the increasing demands in IT industry and for in-depth advanced research.

Textbook
  • Downloadable version: http://dmml.asu.edu/smm/SMM.pdf
  • Print version: https://www.amazon.com/Social-Media-Mining-Reza-Zafarani/dp/1107018854
Useful Lists of Reading Materials

- Books
    - Print version: [https://www.amazon.com/Mining-Social-Web-Facebook-LinkedIn/dp/1449367615](https://www.amazon.com/Mining-Social-Web-Facebook-LinkedIn/dp/1449367615).

- Journals
  - IEEE Transactions on Knowledge and Data Engineering (TKDE)
  - ACM Transactions on Knowledge Discovery from Data (TKDD)
    - ACM Transactions on Information Systems (TOIS)
    - ACM Transactions on Database Systems (TODS)
    - Knowledge and Information Systems (KAIS)
    - DMKD, DKE, TWEB

- Conferences
  - ACM Conference on Knowledge Discovery in Data (KDD)
  - IEEE International Conference on Data Mining (ICDM)
  - International Conference on Web and Social Media (ICWSM)
  - ACM International Conf. on Web Search and Data Mining (WSDM)
  - International World Wide Web Conferences (WWW)
  - IEEE International Conference on Data Engineering (ICDE)
  - Advances in Social Networks Analysis and Mining (ASONAM)
  - ACM Special Interest Group on Information Retrieval (SIGIR)
  - ACM Conf. on Information and Knowledge Management (CIKM)
  - VLDB, SIGMOD, ACMGIS, SDM, AAAI

- Technical Paper Writing
  - Writing Technical Articles, Writing a Technical Report
  - Paper Writing and Presentations
Tentative Course Schedule

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Lecture Topics</th>
<th>Read</th>
<th>Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9/2</td>
<td>Course Introduction; Course project introduction; Introduction of Social Media data;</td>
<td>Ch 1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>9/9</td>
<td>Graph Essentials</td>
<td>Ch 2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>9/16</td>
<td>Network Measures</td>
<td>Ch 3</td>
<td>HW1</td>
</tr>
<tr>
<td>4</td>
<td>9/23</td>
<td>Network Models</td>
<td>Ch 4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>9/30</td>
<td>Data Mining Essentials</td>
<td>Ch 5</td>
<td>Project Proposal</td>
</tr>
<tr>
<td>6</td>
<td>10/7</td>
<td>Information Retrieval Essentials</td>
<td>IR book</td>
<td>HW2</td>
</tr>
<tr>
<td>7</td>
<td>10/14</td>
<td>Midterm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>10/21</td>
<td>Community Analysis</td>
<td>Ch 6</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>10/28</td>
<td>Information Diffusion in Social Media</td>
<td>Ch 7</td>
<td>Project Checkpoint</td>
</tr>
<tr>
<td>10</td>
<td>11/4</td>
<td>Influence and Homophily</td>
<td>Ch 8</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>11/11</td>
<td>Recommendations in Social Media; Behavior Analytics</td>
<td>Ch 9, 10</td>
<td>HW3</td>
</tr>
<tr>
<td>12</td>
<td>11/18</td>
<td>Event Analytics in Social Media</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>11/25</td>
<td>(Thanksgiving break)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>12/2</td>
<td>Project Presentation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>12/9</td>
<td>Project Presentation</td>
<td></td>
<td>Project Report</td>
</tr>
<tr>
<td>16</td>
<td>12/16</td>
<td>Final Exam</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Examinations and Assignments

There are three homework assignments. Homework assignments are due at the start of class. If you have an excused absence from a class, turn in the homework assignment prior to the class session. All assignments must have your name, student ID and course name/number.

The weighting scheme used for grading is: 3 HW Assignments: 20%, including HW1(6%), HW2(7%), and HW3(Research Presentation, 7%); Quiz: 15%; Midterm: 30%; Final Project: 35% (Final Presentation: 10%, Final Report: 25%); and Class Discussion and Participation: 5% (extra bonus). Students are responsible for all material covered in lectures. Examinations will heavily emphasize conceptual understanding of the material.

Important Dates

Last day to add: September 6, 2016
Last day to drop: September 6, 2016 (no tuition penalty)
Final day to drop: September 30, 2016
Last day of selective withdrawal: October 28, 2016
Thanksgiving recess: November 23 - 27, 2016
Last day of classes: December 9, 2016
Final exam: December 16, 2016

Late Submission Policy
Homework assignments and final report must be submitted before the class on the specified due dates (Friday of designated week). A penalty of 30% will be deducted from your score for the first 24-hour period if your submission is late. A penalty of 70% will be deducted from your score for >= 24-hour period. Assignments and report submitted more than 3 days late will not be assessed and will score as a zero (0). Weekend days will be counted. You are encouraged to type your answers.

Course Project Requirement

* Timeline
  o Project proposal: 9/30 (due before class)
  o Project Checkpoint: 10/28 (due before class)
  o Project Presentation: 12/2 and 12/9 (on class)
  o Final Report submission deadline: 12/9 (due before class)

* Project Report
  o IEEE two-column style format, Times New Roman.
  o Font 10.
  o Single space.
  o Each team should submit a separate report:
    - Project proposal: (3 pages)
    - Project checkpoint: (Extended to 6 pages)
    - Final project report (Extended to 8 pages for 2-member team or 10 pages for 3-member team)
  o The project report may contain the following sections:
    - Author (name, student ID, E-mail, and department)
    - Abstract
    - Introduction (including motivation, problem statement, significance of the problem)
    - Related work
    - Proposed approaches
    - System design and implementation (including architecture, datasets, major components, graphic user interface)
    - References

Course Project teams
(to be announced)
**Attendance**

Class attendance is required. Each case of missing class without a proper explanation will cause 5% less from your final numerical grade. If you miss a class, it is your responsibility to find out the material covered in the class. There will absolutely no makeup classes. Only in specific, unavoidable situations students are allowed to excuse absences from class: 1) personal emergencies, including, but not limited to, illness of the student or of a dependent of the student, or death in the family [Require doctor's note]; 2) religious observances that prevent the student from attending class; 3) participation in University-sponsored activities, approved by the appropriate University authority, such as intercollegiate athletic competitions, activities approved by academic units, including artistic performances, academic field trips, and special events connected with coursework; 4) government-required activities, such as military assignments, jury duty, or court appearances; and 5) any other absence that the professor approves.

**Policy on Cheating**

All work is to be done under the provisions of the George Mason University Honor Code: [http://oai.gmu.edu/the-mason-honor-code-2/](http://oai.gmu.edu/the-mason-honor-code-2/). Students can discuss the interpretation of an assignment, however, the actual solution to problems must be one's own. The tenets of the Honor Code will be strictly enforced in this course, and all assignments shall be subject to the stipulations of the Honor Code. Whenever I learn that a student has violated the honor code, I am obligated to report the violation. Cheating in an exam will result in an E grade for the course. Further, the students involved will be referred to the Dean's office for disciplinary action.  

Homework problems are meant to be individual exercises; you must do these by yourself. Any of the following actions will be considered as cheating:

1. A solution which is identical to or nearly identical to the solution submitted by another student in the class.
2. A solution which is identical to or nearly identical to a solution available on the Internet.

Cheating in a homework exercise will result in the following penalty for all the students involved: The homework in which cheating occurred will be assigned a grade of ZERO.

Students who cheat in two or more homework assignments will receive an E grade for the course. The names of such students will also be forwarded to the Dean's office for disciplinary action.

**Disability statement**

If you have a learning or physical difference that may affect your academic work, you will need to furnish appropriate documentation to the Disability Resource Center. If you qualify for accommodation, the DRC staff will give you a form detailing appropriate accommodations for your instructor.
In addition to providing your professors with the appropriate form, please take the initiative to discuss accommodation with them at the beginning of the semester and as needed during the term. Because of the range of learning differences, faculty members need to learn from you the most effective ways to assist you. If you have contacted the Disability Resource Center and are waiting to hear from a counselor, please tell me.