CS 579: Online Social Network Analysis

Introduction

Kai Shu

Spring 2022

Course Objectives

- ## Introducing the basic
  - Social Media Mining (SMM)
- ## Learning the state of the art
  - E.g., graph modeling, adversarial learning, …
- ## Applying what we learn
  - Homework assignments and Projects
- ## Sharing and participating
  - Sample project presentation (learning what others do)
  - Inviting experts to share their research and experience
  - Offering your ideas, information, and feedback

Course Workload and Evaluation

- ## We aim high
  - ### Following the textbook
    - 3 homework assignments and 2 projects
  - ### Advanced Topics (1-2)
    - The advanced research problems that researchers and practitioners are working on
  - ### Exams and Reviews (2)
  - ### Projects and short reports
  - ### How do we grade? Curve or not?
  - ### Late penalty:
    - Exponential penalty: d - #of days, $-\sum 2^{(d-1)}$

- ## Academic integrity
https://web.iit.edu/provost/statement-academic-integrity

Free Textbook

Social Media Mining, Reza Zafarani, Mohammad-Ali Abbasi, Huan Liu, Cambridge University Press, 2014

KDnuggets™ Top Stories, Jun 2, 2018

Featured Story

- **10 More Free Must-Read Books for Machine Learning and Data Science**

Communication Channels and Schedule

- Announcements are made regularly in *Blackboard*
- Emails will be sent out on a need basis
- Weekly Schedule - Please visit Bb for Course Documents
- After-class Blogs (i.e., Discussion Board)
- Office hours (See Bb)
- You're most welcome to share your suggestions and feedback
- We aim to help you succeed in this competitive age with some unique knowledge and skills!

What is this course about?

- Who is this instructor?
  - Kai Shu, and he has a webpage ☺️
  - Assisted by TAs *Canyu Chen* and *Ayush Garg*
- For whom is this course designed?
  - Senior undergraduate and graduate students
  - in particular, for those who want to do research with data science, machine learning, social computing
- Set your goals of attending this course
  - to make you **special** in the age of *data mining* and *machine learning*, or *AI*
- Share your ideas and concerns with us

Social Media

Amazon

Yelp

Twitter

Overview – Dependency Graph
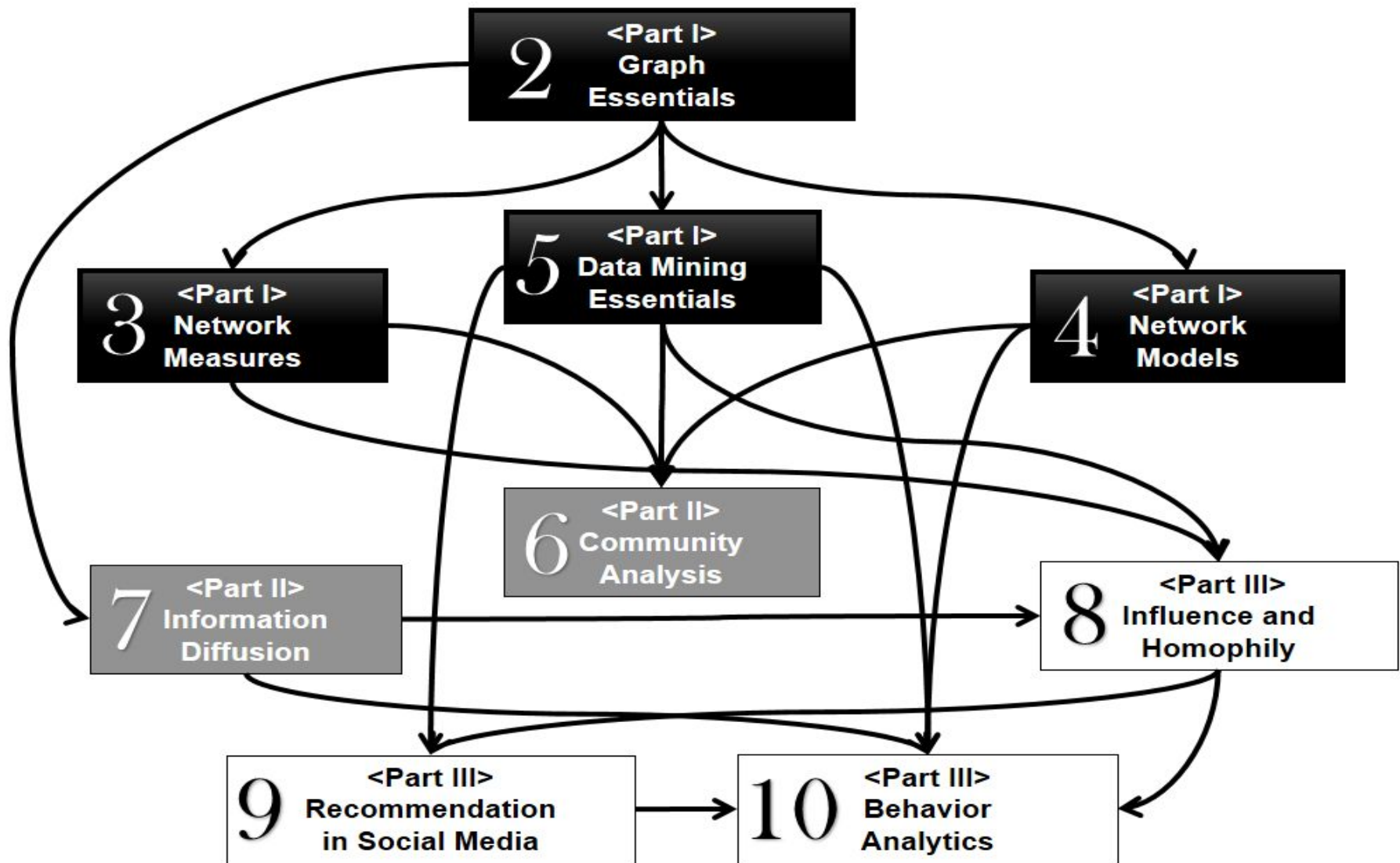
Definition

Social Media is the use of electronic and Internet tools for the purpose of sharing and discussing information and experiences with other human beings in more efficient ways.

Constantly changing …

Social Media Landscape

FredCavazza.net

Social Media: Examples

- A wiki article
- Web reviews and ratings of a popular pizza place in your city
  - E.g., Yelp.com
- An online social network of your professional contacts
  - E.g., Facebook.com, LinkedIn.com
- An iPhone application that informs you where parking is likely available
  - FasPark

Types of Social Media

- Online Social Networking
- Publishing
  - Blogging
  - Wiki
- Micro blogging
- Social News
- Social Bookmarking
- Media Sharing
  - Video Sharing
  - Photo Sharing
  - Podcast Sharing
- Opinion, Review, and Ratings Websites
- Answers
- Entertainment

Online Social Networks

Blogging

Microblogging

Wikis

Social News

Social Bookmarking

Media Sharing

Opinions and Reviews

Answers
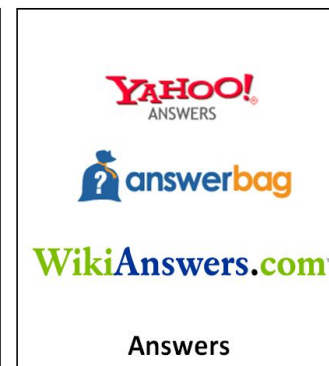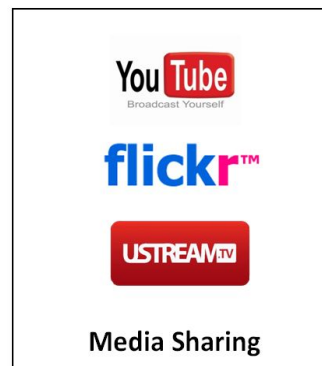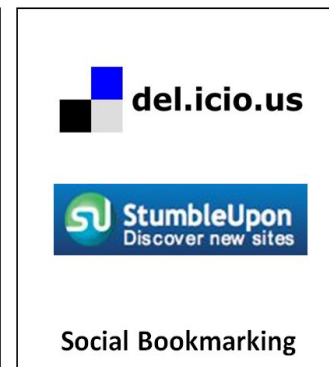
Online Social Networking

Online Social Networks are web-based services that allow individuals and communities to connect with real world friends and acquaintances online

- Interactions
  - Friendship interaction
    - Friends, like, comments, …
  - Media Sharing
  - Sending and receiving messages

- Examples
  - Facebook.com
  - VK.com
  - Snap.com

Blogging

A blog is a journal-like website for users, a.k.a. bloggers, to contribute textual and multimedia content, arranged in reverse chronological order

- Maintained both **individually** or by a **community**

- Usages:
  - Sharing information and opinions with friends and strangers
  - Disseminating subject-specific content
  - Who is the influential

Microblogging

Microblogging is a relatively new phenomenon that can be considered as a counterpart to blogging, but with **limited content**

- Usage
  - communication medium
  - social interaction
  - citizen journalism
- Service Providers:
  - Twitter
  - Google buzz
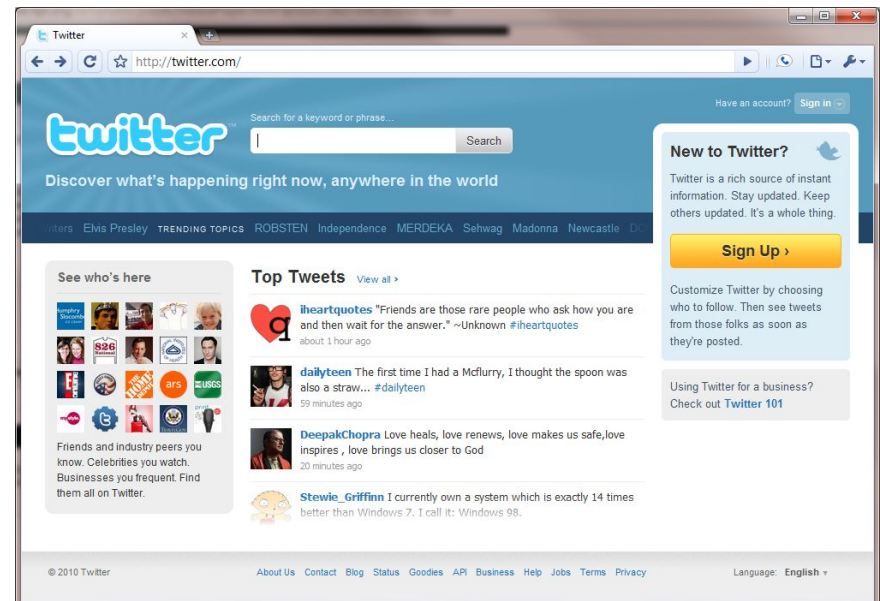
Wiki

A wiki is a collaborative editing environment that allows users to develop Web pages using a simplified markup language

- Wikipedia allows interested individuals to collaboratively develop articles on a variety of subjects.

- Using the wisdom of crowds effectively, it has become a comprehensive repository of information useful to a variety of individuals

Social News

Social News refers to the sharing and selection of news stories and articles by a community of users.

- Users can share articles that they believe would interest the community
- Examples:
  - Digg.com
  - Slashdot
  - Fark
  - Reddit
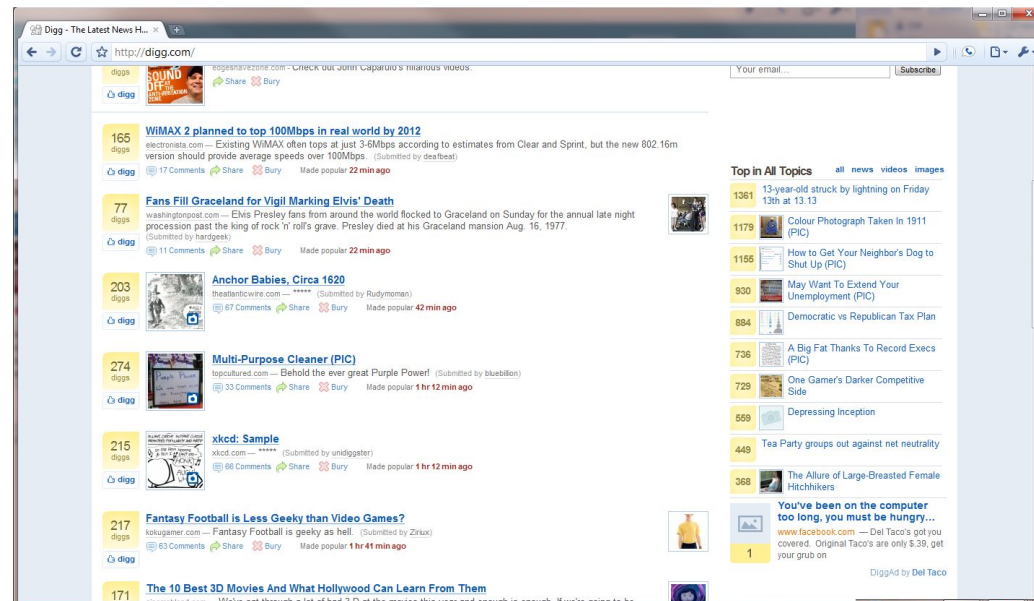
Social Bookmarking

Social Bookmarking sites allow users to bookmark web content for storage, organization and sharing.

- These bookmarks can be tagged with metadata to categorize and provide context to the shared content, allowing users to organize information making it easy to search and identify relevant information.
- Examples
  - Delicious.com
  - StumbleUpon.com
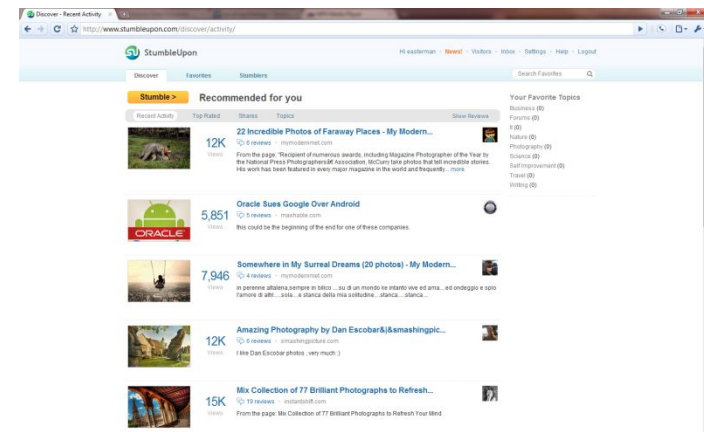
Media Sharing

Media sharing is an umbrella term that refers to the sharing of a variety of media on the web.

Users share such **multimedia content** of possible interest to others

- Examples
  - Video Sharing:
    - YouTube.com
  - Photo Sharing:
    - Flicker.com, picasa.com
  - Document Sharing:
    - Scribd.com, Slideshare.com
  - Livecasting:
    - Justin.tv, Ustream.com
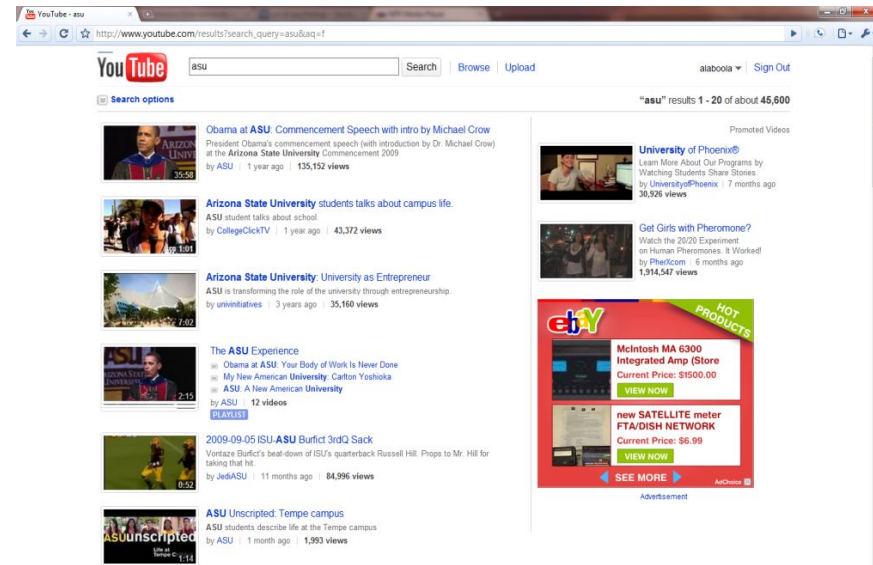
Opinion, Review, and Ratings Websites

Opinion, review, and ratings websites are websites whose primary function is to collect and publish **user-submitted content** in the form of subjective commentary on existing products, services, entertainment, businesses, places, etc.

- Examples
  - Cnet.com
  - Epinions.com
  - yelp.com
  - tripadvisor.com
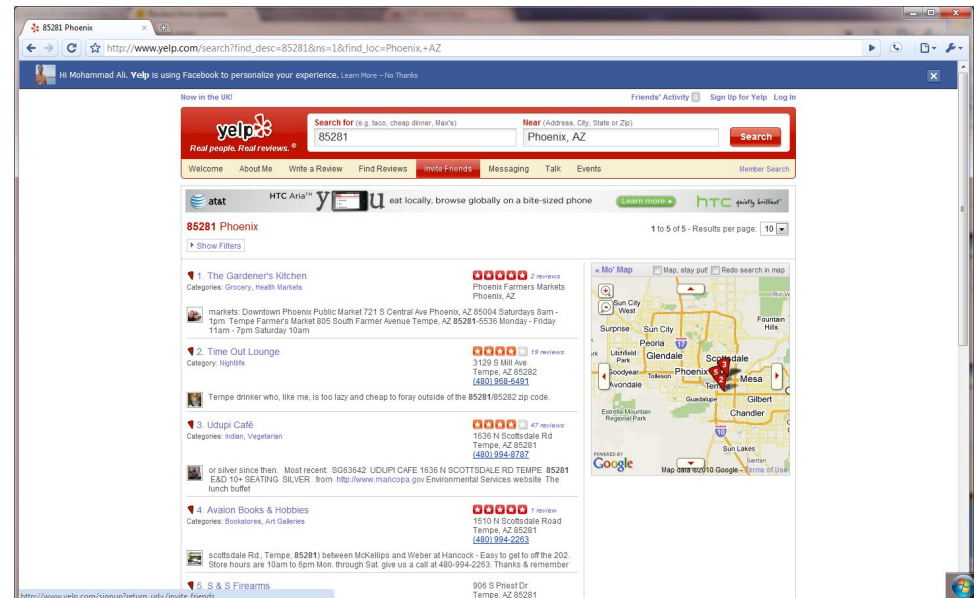
Socially-Provided Answers

In these sites, users who require certain guidance, advice or knowledge can ask **questions**. Other users from the community can answer these questions based on knowledge acquired from previous experiences, personal opinions or from relevant research.

- – Unlike review and opinion sites, which contain self-motivated contribution of opinions, answer sites contain knowledge shared in response to a specific query.
- Examples:
  - – WikiAnswers, Yahoo Answers
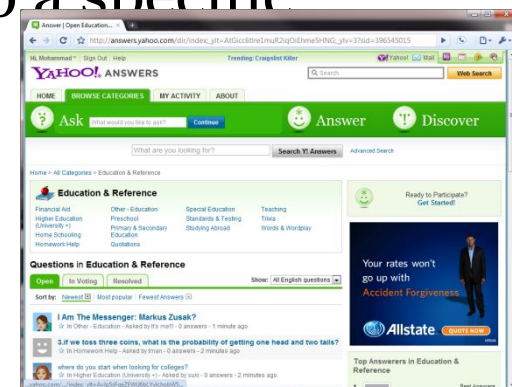
Main Characteristics of Social Media

- ## Participation
  - social media encourages contributions and feedback from everyone who is interested. It blurs the line between media and audience.
- ## Openness
  - most social media services are open to feedback and participation. They encourage voting, comments and the sharing of information. There are rarely any barriers to accessing and making use of content – password-protected content is frowned on.
- ## Conversation
  - whereas traditional media is about "broadcast" (content transmitted or distributed to an audience) social media is better seen as a two-way conversation.
- ## Community
  - social media allows communities to form quickly and communicate effectively. Communities share common interests, such as a love of photography, a political issue or a favorite TV show.
- ## Connectedness
  - most kinds of social media thrive on their connectedness, making use of links to other sites, resources and people.

Social Media Mining *is the process of representing, analyzing, and extracting meaningful patterns from social media data*

Some Challenges in Social Media Mining

1. **Big Data Paradox**
   1. Social media data is big, yet not evenly distributed.
   2. Often little data is available for an individual
2. **Obtaining Sufficient Samples**
   1. Are our samples reliable representatives of the full data?
3. **Noise Removal Fallacy**
   1. Too much removal makes data more sparse
   2. Noise definition is relative and complicated and is task-dependent
4. **Evaluation Dilemma**
   1. When there is no ground truth, how can you evaluate?
5. **Social Data Bias**