

CS 8803 Data Analytics for Well-being: Introduction

Munmun De Choudhury

munmund@gatech.edu

Week 1 | January 11, 2016

Part I: Course Structure and Information

Learning Objectives

- Goal: develop a broad understanding of the emergent field of modeling and analyzing large-scale data for improved well-being and health
- Objectives:
 - Learn about behavioral and social science theories related to this area;
 - Review state-of-the-art data analytic and modeling techniques to understand health behaviors and well-being states;
 - Apply computational and machine learning driven approaches to infer and analyze well-being and health related attributes;
 - Analyze large-scale data (e.g., obtained from social media, mobile health sensors) to understand, model or infer a health phenomenon;
 - Identify the theoretical, empirical, analytical and ethical challenges in employing data analytics and modeling for understanding and modeling well-being and health

Grading

- Off-class discussion, participation in class blog –10% (5% will go toward the 10 reading reflections)
- In-class participation and discussion - 10%
- Class presentation of readings – 5%
- Critique writing –5%
- Take-home assignments – 30% (three, 10% each)
- Term project –40% (project proposal and literature review 5%; development 25%; mid-term presentation 5%; final presentation 5%)
- *There are no exams in this class*

Required Skills

- Technical: any object-oriented/scripting language like Python, Perl, C#;
- Statistics/data mining: preliminary knowledge of working with some data and using some analytical software
- Reading/writing: significant weightage on weekly class readings, mid-term and final term project report

Off-class Participation

- Weightage: 10%
- Breakdown:
 - 10 reflections on papers assigned in different lectures; you can pick the ones of your choice; total 5% (half percent each reflection you submit)
 - 5% toward participation in the course blog/page, most likely a Piazza page

Off-class Participation

- Readings will be available on class website
 - Check the class webpage before each class, as readings are subject to change
- Write short blurbs on Piazza or the class Tumblr blog about the readings assigned for a particular class
 - Blurbs can range from 200-500 words in length
- Blurbs should focus on the following, but not limited to:
 - The novelty in the paper you read i.e., what was the contribution
 - Ways the work in the paper could improve
 - Why the particular method/data used in the paper was appropriate
- Example reflections will be made available on Wednesday

In-class Participation

- Weightage: 10%
- Course will be taught seminar-style, so you are required to participate in the discussion
- Style:
 - Instructor will first cover the material of the assigned readings at a high level
 - Instructor will post questions and intriguing thoughts related to the material
- Student are supposed to engage with these questions and thoughts

Critique Writing

- Weightage: 5%
- Critique writing will involve taking an issue of attention or interest in popular media, relating to the topics covered in class, and developing a critical reflection of it from a technical perspective.
- Due: February 1

Assignments

- Grade weightage: 30% total; three assignments 10% each
- All assignments will be made available three weeks in advance
- Assignments will involve programming and are meant to be individual work

- Assignment I due: February 15
- Assignment II due: March 14
- Assignment III due: April 13

Term Project

- Project Proposal due: February 24
- Mid-term Progress Reports due: March 16
- Final Presentation/demo; Project reports due: May 2

- Group project with ~3 members; you are free to pick your group
- You can pick the topic:
 - Do something related to your research
 - Pick something to replicate from one or more of the class readings
 - Pick a new dataset/health measurement you are curious about
- Need to discuss your project idea with instructor early on in the course

Term Project: What to hand in

- Midterm:
 - If building a tool: design process, and an early prototype
 - If analyzing data: data collection/properties of the data, plan for analysis
 - Report: 4-5 pages, single column format submitted through T-Square
- Final:
 - If building a tool: working prototype, and proposal of its evaluation
 - If analyzing data: choice of the analysis technique, key findings from the analysis (e.g., charts and tables)
 - Report: 8-10 pages, single column format (can include content from midterm for continuity purposes) submitted through T-Square
- Clearly articulate in an extra page individual contribution
- Typically you will not need to submit the code, unless some exception arises
- Okay to use open source tools (in fact encouraged)

Late Policy

- Reading responses are due at 11:59pm on the day before the relevant class meeting.
- Assignments, midterm progress reports and final project reports are due *one hour before class* (i.e., 3:35pm) on the date listed for it.
- 24 hours delay will result in 25% penalty; 48 hours late submissions will incur 50% penalty. Materials submitted past 48 hours will not be accepted, and will be entered as a zero grade.

English as Second Language

- If English is not your first language, you may request to not be graded on your writing for a particular individual assignment/reading reflection
 - This means you won't be penalized for bad writing, but you also won't get credit for good writing. To take advantage of this option, you must mark "ESL" (English as a Second Language) on the first page of your assignment/paper.
- This option is not available for the term project as it is a group assignment.

Academic Integrity

- This class abides by the Georgia Tech Honor Code.
- All assigned work is expected to be individual, except where explicitly indicated otherwise.
- You are encouraged to discuss the assignments with your classmates; however, what you hand in should be your own work.
 - Okay to use open-source software (no need to reinvent the wheel), however do acknowledge!
 - Copying/reusing code from your classmates and friends are not allowed; strict action will be taken if similarities are discovered
 - Copying (textual) content for your assignments and project from other published work (without citing them) is also not allowed, and is considered plagiarism

Help and Resources

- Office hours: 1-2pm Wednesday, or by appointment
- Location: TSRB 341
- Teaching Assistant: TBA

- Class website (including readings):
<http://www.munmund.net/CS8803.html>

- Email announcements will be made over the course page on T-Square

Part II: Introductions

name + program

tech + analytics background

what you want to learn from the class

Topics to be covered

- *Theories* – social and behavioral science, social support
- *Data* – electronic health records, online search logs, social media, wearable sensors
- *Methods* – data modeling and prediction
- *Applications* – health interventions
- *Challenges* – methodological pitfalls, ethics and privacy

Introduction

Why data analytics for well-being?

- Helping doctors and caregivers communicate with patients
- Connecting doctors and patients outside of the formal clinical setting
- Linking patients to patients, or communities identifying with a common well-being concern
- Helping people self-track and stay healthy
- Understanding population-scale ailment challenges, such as the social determinants of health
- Proactive approach toward well-being, instead of reactive approaches
- Beyond traditional medical data...



Watson Health

75





[Google.org home](#)

[Dengue Trends](#)

Flu Trends

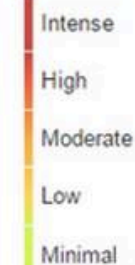
Home

Select country/region ▼

[How does this work?](#)

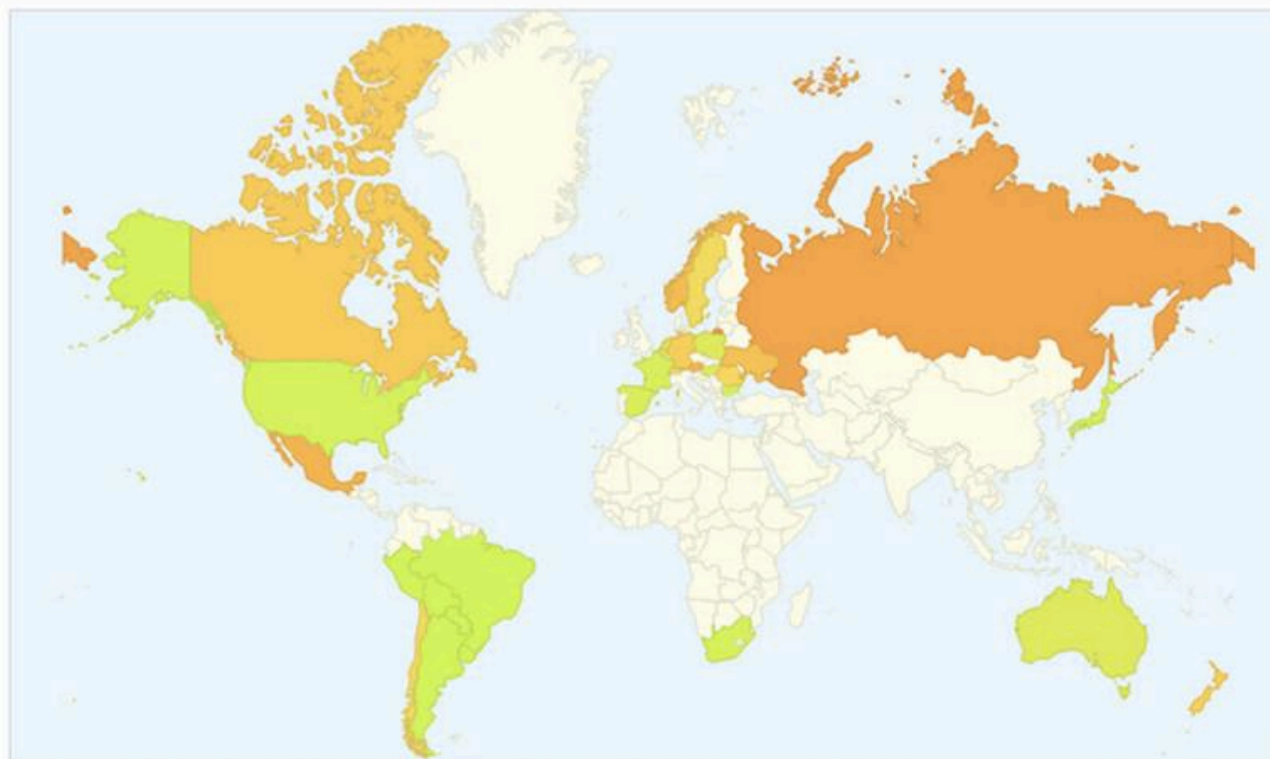
[FAQ](#)

Flu activity



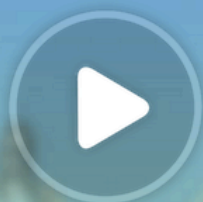
Explore flu trends around the world

We've found that certain search terms are good indicators of flu activity. Google Flu Trends uses aggregated Google search data to estimate flu activity. [Learn more »](#)



[Download world flu activity data](#) - [Animated flu trends for Google Earth](#) - [Compare flu trends across regions in Public Data Explorer](#)

conditions, symptoms, treatments... 🔍



Live better, together!™

Making healthcare better for everyone through sharing, support, and research

[Join now](#)

(it's free!)



Learn from others

Compare treatments, symptoms and experiences with people like you and take control of your health



Connect with people like you

Share your experience, give and get support to improve your life and the lives of others



Track your health

Chart your health over time and contribute to research that can advance medicine for all



STOPSMOKING **hot** new rising controversial top gilded promoted

Want to join? Log in or sign up in seconds. | [English](#)

Welcome new quitters! When you are ready click here to learn how the badge system works

- 1 14 **NEW YEARS QUITTERS! Week #2!** (self.stopsmoking)
 submitted 4 hours ago by FakoPako 10 days
4 comments share
- 2 9 **My advice on cold turkey** (self.stopsmoking)
 submitted 3 hours ago by minusmeminus 10 days
4 comments share
- 3 7 **Day 4 without nicotine** (self.stopsmoking)
 submitted 2 hours ago by Christie3878
2 comments share
- 4 14 **Attempt #5** (self.stopsmoking)
 submitted 6 hours ago by thesavagedetective 1 day
8 comments share
- 5 3 **Starting today** (self.stopsmoking)
 submitted 2 hours ago by use_your_imagination
4 comments share
- 6 4 **72 Hours** (self.stopsmoking)
 submitted 3 hours ago by The_RobberBaron
comment share

search 🔍

username password
 remember me reset password

[Submit a new link](#)

[Submit a new text post](#)

stopsmoking

[subscribe](#) 37,036 Quitters

~60 users here now

This subreddit is a place for redditors to motivate each other to quit smoking. We welcome anyone who wishes to join in by asking or giving advice, sharing stories, or just encouraging someone who is trying to quit.

[How do badges work?](#)

[Self-serve badge instructions](#)

[Request a badge!](#)

[Remove your badge](#)



facebook



Hi Gerald, a friend thinks you might be going through something difficult and asked us to look at your recent post.



Only you can see this. Anything you do there will be kept private.

See Post

Continue



facebook

What would you like to do?

You matter to us, so we want to offer support if you need it. You're not alone—we do this for many people every month. What would you like to do?

Talk to someone

Reach out to a friend or helpline worker.

Get tips and support

Learn how to work through this using some simple tips.

Skip This

Next



search: proana

Everything okay?

If you or someone you know is struggling with an eating disorder, [NEDA](#) is here to help: call 1-800-931-2237 or chat with them online.

If you are experiencing any other type of crisis, consider talking confidentially with a volunteer trained in crisis intervention at www.imalive.org, or anonymously with a trained active listener from [7 Cups of Tea](#).

And, if you could use some inspiration and comfort in your dashboard, go ahead and [follow NEDA on Tumblr](#).

[Go back](#)

[View search results](#)

search: depressed

Everything okay?

If you or someone you know are experiencing any type of crisis, please know there are people who care about you and are here to help. Consider chatting confidentially with a volunteer trained in crisis intervention at www.imalive.org, or anonymously with a trained active listener from [7 Cups of Tea](#).

It might also be nice to fill your dash with inspirational and supportive posts from [TWLOHA](#), [Half of Us](#), the [Lifeline](#), and [Love Is Respect](#).

[Go back](#)

[View search results](#)

Social Media and Well-being

- 18 to 24 year olds are more than 2x as likely than 45 to 54 year olds to use social media for health-related discussions
- 90% of respondents from 18 to 24 years of age said they would trust medical information shared by others on their social media networks
- From a recent study, 54% of patients said they are very comfortable with their providers seeking advice from online communities to better treat their conditions
- 28% of health-related conversations on Facebook are supporting health-related causes, followed by 27% of people commenting about health experiences or updates
- 40% of people polled said information found on social media affects how someone coped with a chronic condition, their view of diet and exercise and their selection of a physician

Challenges

- Information quality and credibility, misinformation
- Privacy risks
- Ethical issues (or the cost of being wrong)
- Helpful support versus promoting damaging behaviors