Natural Language Processing for Health (HLP)
March 2019
Tweet @UPennHLP #HLPMeeting

Graciela Gonzalez-Hernandez
Contact: gragon@upenn.edu
Twitter: @gracielagon

https://healthlanguageprocessing.org
Welcome/Introduction – Graciela Gonzalez-Hernandez

KEYNOTES (15 minutes + 15 minutes for questions each)

Social Media Data for Health Research: Myths and misconceptions
Graciela Gonzalez-Hernandez, PhD, Assistant Professor, Division of Informatics, DBEI, University of Pennsylvania
https://www.dbei.med.upenn.edu/bio/graciela-gonzalez-hernandez-ms-phd

An interpretable natural language processing system for computer-assisted grading of the USMLE step 2 CS examinations
Abeed Sarker, PhD, Research Associate, DBEI, University of Pennsylvania
Social Media for Health Research: Myths and misconceptions

Graciela Gonzalez-Hernandez
@gracielagon
email: gragon@upenn.edu
bio: http://tiny.cc/7gt81y

NLM/NIH grant R01LM011176
PSB 2019 Teri Talk
Myth #1: “NOBODY will say THAT in Social Media!”
"I’ve been diagnosed" + prayers

For the past few weeks I’ve had very low energy. After some blood tests, and conversations with my excellent GP, I’ve been diagnosed with depression. My wife and family know, my church and colleagues know. I figure I’d tweet about it too please keep me in your love and prayers

2:54 AM - 9 Jul 2018

7 Likes
Discovering Cohorts of Pregnant Women From Social Media for Safety Surveillance and Analysis

Abeer Sarker1, PhD; Pramod Chandrashekar2, MS; Arjun Magge2, MS; Haitao Cai1, MA, MS; Ari Klein1, PhD; Graciela Gonzalez2, PhD

1Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States
2Department of Biomedical Informatics, Arizona State University, Scottsdale, AZ, United States

Corresponding Author:
Abeer Sarker, PhD
Department of Biostatistics, Epidemiology and Informatics
Perelman School of Medicine
University of Pennsylvania
Level 4
423 Guardian Drive
Philadelphia, PA, 19104
United States
Phone: 1 6024746203
Email: abeed [at] pennmedicine.upenn.edu

ABSTRACT

Background: Pregnancy exposure registries are the primary sources of information
Myth #2: “You can collect tons of data from social media”
Social Media Mining for Birth Defects Research: A Rule-Based, Bootstrapping Approach to Collecting Data for Rare Health-Related Events on Twitter

Ari Z. Klein, Abeed Sarker, Haitao Cai, Davy Weissenbacher, Graciela Gonzalez-Hernandez

Show more

https://doi.org/10.1016/j.jbi.2018.10.001

Highlights

- Rare health-related events—in this case, birth defects—are reported on Twitter.
- An NLP-based approach was deployed to collect sparse tweets for manual annotation.
The long road to valid data
Finding and validating SM data

- “It’s there” ≠ “you can find it”

- “You can find it” ≠ “a program can find it”

- “A program can find it” ≠ “it is valid”

- Validating SM data requires annotation

- Scalable annotation requires classification
Misconception #1:

“Analyze what you are interested in observing”

aka

“You don’t need controls”
Pharmacoepidemiologic Evaluation of Birth Defects from Health-Related Postings in Social Media During Pregnancy

Abstract

Introduction

Adverse effects of medications taken during pregnancy are traditionally studied through post-marketing pregnancy registries, which have limitations. Social media data may be an alternative data source for pregnancy surveillance studies.

Objective

https://rdcu.be/8tlF
Myth #3:
“Research using social media needs IRB approval as HSR”
Responsible use

- Studies are given an “exempt determination” rather than “IRB approval”.

- Still should abide by responsible conduct of research, ethical and scientific principles

- Still need to use COMMON SENSE

- Although SM data is public, know when to not share, and when to get consent
Misconception #2

“You cannot share SM data”
**Figure:** Representing good science versus misconduct, with a large grey area of Questionable Research Practices (QRP). Data sharing could be QRP or misconduct. Figure adapted from René Custers, VIB (2013).
Solid Social Media Data-based research

- Careful study design
- Well-designed queries/retrieval strategy
- REPRODUCIBLE: data & code sharing
- Hypothesis-driven vs Exploratory
What we have learned

- Social media data can indeed valuable for health research

- It is abundant, but not easy to get

- It is human subjects research but usually receives an “exempt determination”

- Its value depends heavily on good research design and solid scientific practices
Thank you!

Contact: gragon@upenn.edu
Twitter: @gracielagon
bio: http://tiny.cc/7gt81y
Lab: https://healthlanguageprocessing.org
An interpretable NLP system for computer-assisted grading of medical examinations

Improving the grading of the USMLE step 2 CS examinations

Abeed Sarker
Research Associate
Health Language Processing Lab
DBEI, PSOM

03/14/2019
Step 2 CS

- Encounter with standardized patient

- **Patient note (PN)—written in 10 minutes**
  - Data gathering component
    - History
    - Physical examinations
  - Data interpretation
    - Diagnostic tests
    - Diagnoses

- Each examinee has 12 such patient encounters

- Each PN graded by at least 1 physician
Key Essential entries

66 year old male
LLQ abdominal pain x 2 weeks
Worsening
Unrelated to meals
Mild nausea
Subjective fever
Family history of colon cancer
Pain not affected by BMs
Last BM earlier today
Unchanged diet
Negative travel hx
Negative for recent antibiotics
No blood in stool
Alternating constipation/diarrhea for past year

66-year-old man with HTN who presents with 2 weeks of fevers/chills and LLQ abdominal pain. Pain is new for him and does not radiate and is cramping in nature. The intensity waxes and wanes throughout the day but nothing has alleviated pain. Specifically, it is unchanged with Motrin and bowel movements. Also, 2 weeks of subjective fevers/chills with nausea but no vomiting. Has had normal bowel movements with no black or red stools or BRBPR. However, has had a 1.5 year history of periodic constipation and diarrhea. Last colonoscopy at age 50 (was normal). Has not had any FIT tests.

Meds: amlodipine
FHx: mother died of colon cancer at age 72
Diet: fruits, vegetables, lean meat, grains
NLP Challenges

❖ Misspellings
  • Progressivly getting worse.
  • Not alleviated by motrin.
  • Took Motorin with minimal relief.

❖ Abbreviations
  • It was wnl.
  • PMH: HTN, treated with amlodipine.

❖ Unusual sentence structures
  • 66 YO M c/o LLQ pain x 2 weeks, progressively worsening, constant, crampy in nature, started 1/10 in intensity and now became 7/10, not related to food, not alleviated or aggravated with anything.

❖ Inexact references to KE concepts
  • two weeks of increasing lower abd pain (KE: LLQ abdominal pain x 2 weeks)
Scores have to be interpretable at the note level
- A system score must be traceable to matched entries in the Key Essential
- No black box machine learning methods to directly generate a score

Two cases (and KE entries) can be significantly different
- Not possible to use a generalized approach (e.g., Named entity recognition) that can be ported to all the cases
Processing pipeline

- **History Concepts**
- **Physical Examination Concepts**
- **Diagnostic Studies**
- **Diagnoses**

**Key Essentials**

- **Loading, preprocessing and standardizing**
- **Exact and lexicon-based matching**
- **Fuzzy similarity and dynamic thresholding**
- **Supervised concept detection**
- **Annotations**

**Pre-built lexicons**

- **Patient note 1**
  - Hx
  - PE
  - Diagnostic Studies
  - Diagnoses
  - Hx | PE

- **Patient note 2**
  - Hx
  - PE
  - Diagnostic Studies
  - Diagnoses
  - Hx | PE

- **Patient note 3**
  - Hx
  - PE
  - Diagnostic Studies
  - Diagnoses
  - Hx | PE

**History Concepts**

- **Diagnosis**
- **Diagnoses**

**Physical Examination Concepts**

- **PE**

**Diagnostic Studies**

- **Diagnostic Studies**

**Diagnoses**

- **Diagnoses**
- **Hx | PE**

**Concept-level score generation**

- **Set overlap and fuzzy set intersection**

**Key Essentials**

- **Essentials**
Exact and lexicon matching

- **Exact match**: Anorexia (Hx)
  
  - ...Bowel movements have been normal without any frank blood seen. Mild anorexia. No lower urinary tract symptoms...

- **Exact match**: Mild nausea, No weight loss (Hx)
  
  - Ranges from a 3/10-7/10. Has had mild nausea and decrease appetite but no weight loss. Subjective fevers and chills at night but no night sweats.

- **Lexicon match**: Worsening (Hx)
  
  - ... it has gradually been getting worse...
Performance

- Only 1-15% of all the concepts for a case have direct matches
- Lexicons and dictionaries often do not result in observable improvements in score
Fuzzy matching & dynamic thresholding

- **Goal:** Detect expressions that are similar to the KE entries

- **Approach:** A combination of string similarity metrics and thresholding to detect potential inexact matches

- **Examples:**
  - LLQ abdominal pain x 2 weeks > llq abdominal pain started 2 weeks (0.86)
  - No blood in stool > no bloody stools (0.97)
  - Abdominal tenderness LLQ > tender to palpation over llq (0.81)
  - Alternating constipation/diarrhea for past year > alternating diarrhea and constipation for which he takes (0.85)
  - Unchanged diet > no changes in his diet (0.67)
  - Family history of colon cancer > fh mother died of colon cance (0.66)
  - Pain not affected by BMs > bowl movement freq not assoc with abd pain (0.47)

Dynamic Threshold = \( T_i - (k \times KE_{\text{entry\_tokens}}) \)
Incorporating human supervision: Annotation

- KE Entries

66 year old male
LLQ abdominal pain x 2 weeks
Worsening
Unrelated to meals
Mild nausea
Subjective fever
Family history of colon cancer
Pain not affected by BMs
Last BM earlier today
Unchanged diet
Negative travel hx
Negative for recent antibiotics
No blood in stool
Alternating constipation/diarrhea for past year

66-year old man with HTN who presents with 2 weeks of fevers/chills and LLQ abdominal pain. Pain is new for him and does not radiate and is cramping in nature. The intensity waxes and wanes throughout day but nothing has alleviated pain. Specifically, it is unchanged with Motrin and bowel movements. Also 2 weeks of subjective fevers/chills with nausea but no vomiting. Has had normal bowel movements with no black or red stools or BRBPR. However, has had a 1.5 year history of periodic constipation and diarrhea. Last colonoscopy at age 50 (was normal). Has not had any FIT tests.

Meds: amlodipine

FHx: mother died of colon cancer at age 72

Diet: fruits, vegetables, lean meat, grains
Annotation: increased search space

KE entry:
- increased thirst

Annotated entries:
- polydyspsia
- thirsty
- increased water intake
- POLYDIPSIA
- concurrent increase in fluid intake
- HAVING TO DRINK MORE WATER
Other methods and features

- Set overlap and fuzzy overlap to detect concepts spread out over long sentences
  - Even if some of the terms are spelled incorrectly

PN text: “No headache, nausea, subjective fever, weakness or family history of diabetes”

KE text: “No history of diabetes”

- Patterns for detecting age and gender mentions

- Other optimizations for speed, increasing annotation coverage and customizability
Performance

Mean F<sub>1</sub>-score = 0.89
Contribution of each set of functions

![Graph showing F-score and average number of matches per encounter for different sets of functions.](image-url)
Error analysis

**False Positives**
- System prediction wrong and/or similarity threshold too low (with annotation)
- Annotation issue/mismatch
- Incorrect partial match/detected partially correct answer
- Matches incorrect KE
- Incorrect pattern

**False Negatives**
- Long expressions/threshold too high
- Creative or unseen expression (ambiguous)
- Undetected negation
- Very different texts
- Incorrect annotation
- Other
Comparison of system scores with actual scores