People on Drugs: Credibility of User Statements in Health Forums

Subhabrata Mukherjee $^1$ Gerhard Weikum $^1$
Cristian Danescu-Niculescu-Mizil $^2$

$^1$Max Planck Institute for Informatics
$^2$Max Planck Institute for Software Systems

KDD 2014

August 25, 2014
Motivation: Internet as a healthcare resource

59% of US population use internet for health information

[Pew Research Center Report, 2013]

Half of US physicians rely on online resources

[IMS Health Report, 2014]

This work:

Credibility of user-generated online health information
Motivation: Internet as a healthcare resource

59% of US population use internet for health information
[Pew Research Center Report, 2013]

Half of US physicians rely on online resources
[IMS Health Report, 2014]

This work: 
Credibility of user-generated online health information
“My girlfriend always gets a bad dry skin, rash on her upper arm, cheeks, and shoulders when she is on [Depo]. . . . ”

“I have had no side effects from [Depo] (except . . . ), but otherwise no rashes. She should see her gyno. She may be allergic to something”
“My girlfriend always gets a bad dry skin, rash on her upper arm, cheeks, and shoulders when she is on [Depo]. . . .”

“I have had no side effects from [Depo] (except ... ), but otherwise no rashes. She should see her gyno. She may be allergic to something”
Our Intuition

Users, language and credibility influence each other.

I took a cocktail of meds. Xanax gave me hallucinations and a demonic feel. Xanax and Prozac are known to cause drowsiness. Xanax made me dizzy and sleepless.

Trustworthy users write credible posts. Agree with each other on credible statements.
I took a cocktail of meds. Xanax gave me hallucinations and a demonic feel. Xanax and Prozac are known to cause drowsiness. Xanax made me dizzy and sleepless.
“I heard Xanax can have pretty bad side-effects. You may have peeling of skin, and apparently some friend of mine told me you can develop ulcers in the lips also. If you take this medicine for a long time then you would probably develop a lot of other physical problems. Which of these did you experience?”

Usage of modals, indefinite determiner, conditional, probabilistic adverb, question particle, etc.
“I heard Xanax can have pretty bad side-effects. You may have peeling of skin, and apparently some friend of mine told me you can develop ulcers in the lips also. If you take this medicine for a long time then you would probably develop a lot of other physical problems. Which of these did you experience?”

Usage of modals, indefinite determiner, conditional, probabilistic adverb, question particle, etc.
"I heard Xanax can have pretty bad side-effects. You may have peeling of skin, and apparently some friend of mine told me you can develop ulcers in the lips also. If you take this medicine for a long time then you would probably develop a lot of other physical problems. Which of these did you experience?"

Usage of modals, indefinite determiner, conditional, probabilistic adverb, question particle, etc.
“I heard Xanax can have pretty bad side-effects. You may have peeling of skin, and apparently some friend of mine told me you can develop ulcers in the lips also. If you take this medicine for a long time then you would probably develop a lot of other physical problems. Which of these did you experience?”

Usage of modals, indefinite determiner, conditional, probabilistic adverb, question particle, etc.
“I heard Xanax can have pretty bad side-effects. You may have peeling of skin, and apparently some friend of mine told me you can develop ulcers in the lips also. If you take this medicine for a long time then you would probably develop a lot of other physical problems. Which of these did you experience ?”

Usage of modals, indefinite determiner, conditional, probabilistic adverb, question particle, etc.
“Depo is very dangerous as a birth control and has too many long term side-effects like reducing bone density. Hence, I will never recommend anyone using this as a birth control. Some women tolerate it well but those are the minority. Most women have horrible long lasting side-effects from it.”

Uses *inferential conjunction*, modal, definite determiners, *etc.*
“Depo is very dangerous as a birth control and has too many long term side-effects like reducing bone density. Hence, I will never recommend anyone using this as a birth control. Some women tolerate it well but those are the minority. Most women have horrible long lasting side-effects from it.”

Uses inferential conjunction, modal, definite determiners, etc.
“Depo is very dangerous as a birth control and has too many long term side-effects like reducing bone density. Hence, I will never recommend anyone using this as a birth control. Some women tolerate it well but those are the minority. Most women have horrible long lasting side-effects from it.”

Uses inferential conjunction, modal, definite determiners, etc.
“I started Cymbalta, but now I’m having a **panic attack** or an allergic reaction. I have a hardcore burning sensation in my chest and warm sensations all over. It’s like my body can’t decide whether it wants to be cold or hot. I feel if I close my eyes I’ll lose control, go **crazy** and pass out.”
Our Intuition

I took a cocktail of meds. Xanax gave me hallucinations and a demonic feel.

Xanax and Prozac are known to cause drowsiness.

Xanax made me dizzy and sleepless.

User Trustworthiness

Language Objectivity

Statement Credibility
User Features

▶ User demographic features like age, gender, location

▶ Engagement features like number of posts, questions, answers, thanks

▶ User post properties like avg. post length
I took a cocktail of meds. **Xanax** gave me hallucinations and a demonic feel. **Xanax** and Prozac are known to cause drowsiness. **Xanax** made me dizzy and sleepless.

This is what we want
I took a cocktail of 
meds. **Xanax** gave 
me **hallucinations** 
and a demonic feel.

**Xanax** and **Prozac** 
are known to 
cause **drowsiness**.

**Xanax** made me 
dizzy and sleepless.

Predict the *most likely label assignment* of statements
Semi Supervised Learning

Protects against users conveying misinformation using confident and objective language

I took a cocktail of meds. Xanax gave me hallucinations and a demonic feel.

Xanax and Prozac are known to cause drowsiness.

Xanax made me dizzy and sleepless.

Observed Features

User Trustworthiness

u1

u2

u3

Language Objectivity

p1

p2

p3

Statement Credibility

s1

s2

s3?

CRF

Labels ?

Expert stated side-effects of drugs from MayoClinic portal
Semi-Supervised CRF (Sketch)

User Trustworthiness

Language Objectivity

Statement Credibility
Semi-Supervised CRF (Sketch)

User Trustworthiness
- $u_1$
- $u_2$
- $u_3$

Language Objectivity
- $p_1$
- $p_2$
- $p_3$

Statement Credibility
- $s_1$
- $s_2$?

- True
- False
- Unknown
Semi-Supervised CRF (Sketch)

User Trustworthiness

Language Objectivity

Statement Credibility

Depo \rightarrow \text{dry skin}
1. Estimate user trustworthiness:

\[ t_k = \frac{\sum_{i} \mathbb{1}_{S_{i,k}=\text{True}}}{|S_k|} \]
1. Estimate user trustworthiness:

\[ t_k = \frac{\sum_i \mathbb{1}_{s_{i,k}=\text{True}}}{|S_k|} \]

- User Trustworthiness:
  - \( u_1 \)
  - 0.5
  - \( u_2 \)
  - \( u_3 \)
  - 1

- Language Objectivity:
  - \( p_1 \)
  - \( p_2 \)
  - \( p_3 \)

- Statement Credibility:
  - \( s_1 \)
  - \( s_2 \)
2. E-Step: Estimate label of unknown statements by Gibbs' sampling:

\[ Pr(S_i^U | P, U, S^L; W) \propto \prod_{\nu \in C} t_k \times \phi_\nu(S^*_\nu, p_j, u_k; W) \]
2. E-Step: Estimate label of unknown statements by Gibbs' sampling:

\[ Pr(S_i^U | P, U, S^L ; W) \propto \prod_{\nu \in C} t_k \times \phi_\nu(S_\nu^*, p_j, u_k ; W) \]
3. M-Step: Maximize log-likelihood to estimate feature weights using Trust Region Newton:

\[ W^{(\nu+1)} = \arg\max_W \sum_{SU} q(S^U) \log Pr(S^L, S^U | P, U; W') \]
4. Re-Estimate user trustworthiness: 

\[ t_k = \frac{\sum_i \mathbb{1}_{S_{i,k}=\text{True}}}{|S_k|} \]
4. Re-Estimate user trustworthiness: 

$$t_k = \frac{\sum_i 1_{S_{i,k}=\text{True}}}{|S_k|}$$
4. Re-Estimate user trustworthiness: \( t_k = \frac{\sum_i \mathbb{1}_{S_{i,k} = \text{True}}}{|S_k|} \)

5. Apply \textit{E-Step} and \textit{M-Step} until convergence
Dataset

Healthboards.com community (www.healthboards.com) with 850,000 registered users and 4.5 million messages
▶ We sampled 15,000 users with 2.8 million messages

Expert labels about drugs from MayoClinic portal
▶ 2172 drugs categorized in 837 drug families
▶ 6 widely used drugs used for experimentation
Dataset

Healthboards.com community (www.healthboards.com) with 850,000 registered users and 4.5 million messages
▶ We sampled 15,000 users with 2.8 million messages

Expert labels about drugs from MayoClinic portal
▶ 2172 drugs categorized in 837 drug families
▶ 6 widely used drugs used for experimentation
Drug Statistics

Data available at: http://www.mpi-inf.mpg.de/impact/peopleondrugs/

<table>
<thead>
<tr>
<th>Drugs</th>
<th>Treatment For</th>
<th># Users</th>
<th># Posts</th>
</tr>
</thead>
<tbody>
<tr>
<td>alprazolam</td>
<td>anxiety, depression, panic disorder</td>
<td>2.8K</td>
<td>21K</td>
</tr>
<tr>
<td>ibuprofen</td>
<td>pain, symptoms of arthritis</td>
<td>5.7K</td>
<td>15K</td>
</tr>
<tr>
<td>omeprazole</td>
<td>acidity in stomach and ulcers</td>
<td>1K</td>
<td>4K</td>
</tr>
<tr>
<td>metformin</td>
<td>high blood sugar, diabetes</td>
<td>.8K</td>
<td>3.6K</td>
</tr>
<tr>
<td>levothyroxine</td>
<td>hypothyroidism</td>
<td>.4K</td>
<td>2.4K</td>
</tr>
<tr>
<td>metronidazole</td>
<td>bacterial infection</td>
<td>.5K</td>
<td>1.6K</td>
</tr>
</tbody>
</table>
Baselines

- Frequency of statements

- SVM Classification
  - Feature vector for each statement using all our features

- SVM Classification with Distant Supervision
  - Each user, post and statement instance constitutes a feature vector
  - Aggregate labels of all such instances for a statement by majority voting
Accuracy Comparison

![Bar Chart]:
- **Alprazolam**
- **Metronidazole**
- **Omeprazole**
- **Levothyroxine**
- **Metformin**
- **Ibuprofen**

Accuracy

- **Freq.**
- **SVM**
- **SVM Dist. SV.**
- **CRF**
Use-Case: Following Trustworthy Users

What users should I follow to get information on drug X?

Baseline: Rank users based on #thanks from community
Use-Case: Following Trustworthy Users

Compare with human annotations

![Bar chart showing NDCG for Alprazolam and Levothyroxine for Thanks and CRF]
Conclusions

Proposed a probabilistic graphical model to jointly learn *user trustworthiness, statement credibility and language use*

- To extract side-effects of drugs from communities
- Identify expert users

Provides a framework to incorporate richer linguistic (e.g., bias, discourse) and user (e.g., perspective, expertise) features
Thank you