Identifying Controversial Topics in Large-scale Social Media Data

Master’s Thesis

Olaoluwa Anifowose
Fakultät Medien
Bauhaus-Universität Weimar
31.03.2016

Supervised by:
Prof. Dr. Benno Stein
Prof. Dr. Sven Bertel

Advisors:
Dr. Henning Wachsmuth
Michael Völske
Outline

1. Introduction
2. Motivation
3. Objectives
4. Reddit
5. Approach
6. Experiments and Evaluation
Introduction
Controversies and the Web

Controversy
A state of prolonged public dispute or debate, usually concerning a matter of conflicting opinion or point of view.

Controversies on the web
- The web: a platform to express opinions or point of views
- A place to identify controversies.
Motivation

Effects of Controversies
- Facilitates productive debates
- Provides balanced view about a topic
- Can lead to tension and conflict

Why detect controversies
- Avoid or detect tension and conflict
- Used by search engines in meeting users information need
Motivation

Manual Controversy Detection Challenges

- The size of the web
- The diversity of the web
- The dynamic nature of the web
- The context and scope of controversy
Motivation

Automatic Controversy Detection: Related Work

- Controversy Ranked Models [Vuong et al 2008]
  - Objective: Detect if a page is controversial
Motivation

Automatic Controversy Detection: Related Work

- **Controversy Ranked Models [Vuong et al 2008]**
  - Objective: Detect if a page is controversial

- **Regression Machine Learning [Popescu et al 2010]**
  - Domain: Twitter
  - Objective: Ranking of how controversial a discussion is
Motivation
Automatic Controversy Detection: Related Work

- **Controversy Ranked Models [Vuong et al 2008]**
  - Objective: Detect if a page is controversial

- **Regression Machine Learning [Popescu et al 2010]**
  - Domain: Twitter
  - Objective: Ranking of how controversial a discussion is

- **Fine grained [Bykau et al 2015]**
  - Objective: What, Where, When, Who
Objectives

Research Questions

- How easily identifiable are controversies in a large domain?
- How early can these controversies be detected?
- How effective are the known controversial measures in detecting controversies?
Reddit

How I see the US presidential election as a nerd.

submitted 10 days ago by bawbrosss
4449 comments share

Top 200 comments show 500

[-] Joal0503 3306 points 10 days ago

Ted cruz as "The Penguin"

[-] Anselan 3388 points 10 days ago

I was about to Photoshop that up, and I just decided to do a google search: http://i.imgur.com/2pDjZcL.jpg

[-] DMann420 1141 points 10 days ago

Well that's convenient.

[-] wnbaloll 526 points 10 days ago

Internet has done it again

[-] d_haven 420 points 10 days ago

Great job everyone! Let's go home.
Reddit
Reddit Dataset

196.5M submissions & 1.7B comments
July, 2006 to May, 2015

370M Authors
Reddit

Reddit Representation

- A page \( p = (s, c, t) \). Where
  - \( s \) is the page submission
  - \( c \), the comments associated with the submission
  - \( t \), the time the submission was created

- A page is controversial if it has at least one controversial comment
  - Using the the *controversiality* field of comments metadata

- Equal number of controversial and non-controversial pages
  - Balanced across subreddit and time
### Reddit Dataset Balancing

<table>
<thead>
<tr>
<th>Subreddit</th>
<th>Time</th>
<th>#Non-Controversial Pages</th>
<th>#Controversial Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>sub₁</td>
<td>m₁_y₁</td>
<td>n₁₁</td>
<td>n₁₁</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>m₁_y₁</td>
<td>n₁₁</td>
<td>n₁₁</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>subᵢ</td>
<td>mᵢ_yᵢ</td>
<td>nᵢᵢ</td>
<td>nᵢᵢ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mᵢ_yᵢ</td>
<td>nᵢᵢ</td>
<td>nᵢᵢ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Approach

Problem Representation

Supervised Learning Problem

- Label data as controversial or not
Approach

Problem Representation

Supervised Learning Problem

- Label data as controversial or not
- Learn from examples from the past
  - Training set
Approach

Problem Representation

Supervised Learning Problem

- Label data as controversial or not
- Learn from examples from the past
  - Training set
Approach
Problem Representation

Supervised Learning Problem
- Label data as controversial or not
- Learn from examples from the past
  - Training set
- Make predictions on recent data
  - Test set
Approach

Problem Representation

Supervised Learning Problem
- Label data as controversial or not
- Learn from examples from the past
  - Training set
- Make predictions on recent data
  - Test set
Approach

Feature Engineering

- Structural Features

- Linguistic Features

- Sentiment Features

- Age Dependent Features
Approach
Feature Engineering

- **Structural Features**
  - Features derived from the composition of the data
  - Mainly from the metadata of the dataset
  - Examples: number of comments, depth of comments, average timestamp of comments etc.

- **Linguistic Features**

- **Sentiment Features**

- **Age Dependent Features**
Approach
Feature Engineering

- **Structural Features**
  - Features derived from the composition of the data
  - Mainly from the metadata of the dataset
  - Examples: number of comments, depth of comments, average timestamp of comments etc.

- **Linguistic Features**
  - Describes the text content of the entire page.
  - Examples: percentage of nouns, percentage of verbs etc.

- **Sentiment Features**

- **Age Dependent Features**

Olaoluwa Anifowose (Bauhaus Universität Weimar) - Identifying Controversial Topics in Large-scale Social Media Data - Master’s Thesis Defense, 2016
Approach

Feature Engineering

- **Structural Features**
  - Features derived from the composition of the data
  - Mainly from the metadata of the dataset
  - Examples: number of comments, depth of comments, average timestamp of comments etc.

- **Linguistic Features**
  - Describes the text content of the entire page.
  - Examples: percentage of nouns, percentage of verbs etc.

- **Sentiment Features**
  - Reveals the views and opinions expressed by the authors
  - Examples: fraction of positive words, fraction of negative words etc.

- **Age Dependent Features**
Approach

Feature Engineering

- **Structural Features**
  - Features derived from the composition of the data
  - Mainly from the metadata of the dataset
  - Examples: number of comments, depth of comments, average timestamp of comments etc.

- **Linguistic Features**
  - Describes the text content of the entire page.
  - Examples: percentage of nouns, percentage of verbs etc.

- **Sentiment Features**
  - Reveals the views and opinions expressed by the authors
  - Examples: fraction of positive words, fraction of negative words etc.

- **Age Dependent Features**
  - Takes advantage of the age of the page
  - Examples: comments per age etc
Experiments and Evaluation

Experiments Overview

- **Time Independent Experiments**
  - **Training Set:** All comments of the page
  - **Test Set:** All comments of the page
Experiments and Evaluation

Experiments Overview

- **Time Independent Experiments**
  - **Training Set:** All comments of the page
  - **Test Set:** All comments of the page

- **Time Dependent Experiments**
  - **Training Set:** Only comments of the page at specified age of the page
  - **Test Set:** Only comments of the page at specified age of the page
Experiments and Evaluation

Experiments Overview

- **Time Independent Experiments**
  - **Training Set**: All comments of the page
  - **Test Set**: All comments of the page

- **Time Dependent Experiments**
  - **Training Set**: Only comments of the page at specified age of the page
  - **Test Set**: Only comments of the page at specified age of the page

- **Hybrid Experiments**
  - **Training Set**: All comments of the page
  - **Test Set**: Only comments of the page at specified age of the page
Experiments and Evaluation

Evaluation Metrics

- **Precision** = \( \frac{TP}{TP + FP} \)
- **Recall** = \( \frac{TP}{TP + FN} \)
- **F-Score** = \( 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \)
Experiments and Evaluation

Experiment Results

Table: Time Independent Experiment Test Data Result

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Correctly Predicted</td>
<td>75.4</td>
</tr>
<tr>
<td>Precision</td>
<td>0.84</td>
</tr>
<tr>
<td>Recall</td>
<td>0.63</td>
</tr>
<tr>
<td>F-Score</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Objective

- How easily identifiable are controversies in a large domain?
Experiments and Evaluation

Experiment Results

Objective

- How early can controversies be detected?
Objective

- Effect of incomplete information in classification.
Experiments and Evaluation

Experiment Results

Table: Features Information Gain Ranking of highest ranked features in each Feature Family

<table>
<thead>
<tr>
<th>Features</th>
<th>Family</th>
<th>Info Gain</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>numOfComments</td>
<td>Structural</td>
<td>0.422</td>
<td>1st</td>
</tr>
<tr>
<td>maximumNumOfComments-InTimeSequences</td>
<td>Age Dependent</td>
<td>0.399</td>
<td>2nd</td>
</tr>
<tr>
<td>controversialMix</td>
<td>Sentiment</td>
<td>0.371</td>
<td>4th</td>
</tr>
<tr>
<td>percentageOfQuestions</td>
<td>Linguistic</td>
<td>0.325</td>
<td>9th</td>
</tr>
</tbody>
</table>

Objective

- How effective are the known controversial measures in detecting controversial topics?
Experiments and Evaluation

Experiment Results

Table: Features Information Gain Ranking of lowest ranked features in each Feature Family

<table>
<thead>
<tr>
<th>Features</th>
<th>Family</th>
<th>Info Gain</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>numOfGildes</td>
<td>Structural</td>
<td>0.001</td>
<td>32nd</td>
</tr>
<tr>
<td>percentageOfSarcastics</td>
<td>Linguistic</td>
<td>0.026</td>
<td>28th</td>
</tr>
<tr>
<td>commentsPerAge</td>
<td>Age Dependent</td>
<td>0.063</td>
<td>27th</td>
</tr>
<tr>
<td>contradictionScore</td>
<td>Sentiment</td>
<td>0.216</td>
<td>18th</td>
</tr>
</tbody>
</table>

Objective

- How effective are the known controversy detection measures
Summary

Contributions

• A balanced social media dataset that can be used for research.
Summary

Contributions

- A balanced social media dataset that can be used for research.
- Knowledge about how early controversies can be detected.
  - Controversy can be detected within 0.3 days (7 hours 12 minutes)
Summary

Contributions

- A balanced social media dataset that can be used for research.
- Knowledge about how early controversies can be detected.
  - Controversy can be detected within 0.3 days (7 hours 12 minutes)
- Effect of incomplete information in classification.
  - Lower percentage of page correctly classified for the first two days
Summary

Contributions

- A balanced social media dataset that can be used for research.
- Knowledge about how early controversies can be detected.
  - Controversy can be detected within 0.3 days (7 hours 12 minutes)
- Effect of incomplete information in classification.
  - Lower percentage of page correctly classified for the first two days
- Performance of known controversy detection measures.
  - Sentiment Features: a good measure of controversy
  - Age Dependent Features: a good measure for data with incomplete information
Summary

Future Work

- Effective way of topic extraction
- Feature Engineering: Development of more features
Thank You