SCREENING FOR PTSD USING VERBAL FEATURES IN SELF-NARRATIVES
A TEXT MINING APPROACH
02-07-2014

Sytske Wiegersma
Qiwei He
Contents

- Posttraumatic Stress Disorder (PTSD)
- Study 1 (finished): Text mining for PTSD screening
- Study 2 (starting up): Text/audio mining for PTSD treatment
- Conclusion
Posttraumatic Stress Disorder (PTSD)

Definition:

“Posttraumatic Stress Disorder (PTSD) is an anxiety disorder that can develop after exposure to one or more traumatic events that threatened or caused great physical harm”. (Source: Brunet, Akerib & Birmes, 2007)
Posttraumatic Stress Disorder (PTSD)

Characteristics:
- Traumatic event is not successfully processed
- Symptoms do not subside over time
- Patient suffers psychologically, physically, emotionally and socially
- Can be prevented by early diagnosis and treatment
Diagnostic criteria

Core diagnostic criteria – DSM-5, APA, 2013

Criterion A - Stressor

Criterion B - Intrusion

Criterion C - Avoidance

Criterion D – Negative alterations in cognitions and mood

Criterion E – Alterations in arousal and reactivity
Study 1: Text mining for PTSD screening

- Currently used instruments:
  - Clinical diagnosis: Structured Interview for PTSD
  - Prognostic screening: Questionnaires on risk factors for PTSD

- Research trauma narratives - PTSD symptom severity:
  - Narrative content (Pennebaker, 1993)
  - Narrative organization and cohesiveness (Foa et al., 1995)

- Computerized text classification model for PTSD screening (Qiwei He, 2013)

Accurate, but:
- False positives
- Time consuming
- Socially desirable answers

Call for innovative screening strategies.

Screening instrument based on trauma narratives.
Text mining

“Text mining seeks to extract useful information from document collections through the identification and exploration of patterns among unstructured textual data”. (Source: Feldman & Sanger, 2007)

- No pre-defined data model
- Not organized
- Typically text-heavy

Structured Data

- Highly organized
- Readily searchable
- Easy to identify and process

Machine learning: Techniques and algorithms to make predictions based on known properties learned from training data.

Data Mining

UNIVERSITY OF TWENTE.
Machine Learning Process

- Preparation
  - Sample Collection 300
  - Train / Test Set
  - Preprocessing

- Training
  - Input Train Set
  - Keywords Extraction
  - Chi-Square Selection
  - Weights Allocation
  - Output Weighted Stems

- Testing
  - Input Test Set
  - Cue-Stems Scanning
  - Classification Rule
  - Output PTSD +/-
Dataset

- **Sample Source:**
  - 300 English narratives from online forum for patients with mental diseases

- **Sample Requirements:**
  - First episode description
  - Content: description of traumatic event and symptoms
  - Official diagnosis (PTSD/Non-PTSD) via structured interview by psychiatrists

Total data set (300 stories)

PTSD + (150 stories)
- 50
- 100

PTSD - (150 stories)
- 100
- 50

Train Set (200 stories)

Test Set (100 stories)
Preprocessing

- **Stop Word List:**
  170-stop-word list was used to filter out words with little meaning, including
  (a) 128 words in English stop list embedded in Python (e.g., “I”, “is”)
  (b) 10 common used punctuations (e.g., “.”, “,”)
  (c) 32 abbreviations (e.g., “isn’t”, “I’m”)

- **Stemming:** *Porter stemming algorithm* (Porter, 1980) to remove plurals and suffixes

- **Text representations:**
  - Unigrams: one single word
  - Bigrams: two subsequent words
  - Trigrams: three subsequent words
  - Unigrams + Bigrams: combination of single and two subsequent words
  - Mix N-grams: Combination of Unigrams + Bigrams + Trigrams
Key-word extraction

Chi-Square Selection Algorithm

\[ X^2 = \frac{N(O_{11}O_{22} - O_{12}O_{21})^2}{(O_{11} + O_{12})(O_{11} + O_{21})(O_{12} + O_{22})(O_{21} + O_{22})} \]

- Tests whether occurrence of word and occurrence in specific class are dependent.
- The higher Chi-Square score, the more informative the stem.
- Rank the stems in decreasing order, select the top stems as robust classifiers.

<table>
<thead>
<tr>
<th>Stem</th>
<th>PTSD+ C₁</th>
<th>PTSD- C₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stem1</td>
<td>45</td>
<td>1</td>
</tr>
<tr>
<td>Stem2</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>Stem3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Stem200</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Stem3860</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

\[ \sum_{k=1}^{k} n_k m_k = \text{len}(C_1) \quad \text{len}(C_2) \]
Weights allocation

Three machine learning algorithms were compared:

Accuracy Rate of Three Models in Cross Validation

- **Standard classifiers:**
  - Decision Tree
  - Naïve Bayes

- **Self-designed classifier:**
  - Product Score model
### Results

#### Striking:
- Accuracy rate decreases when using bigrams and trigrams.
- Accuracy rate is not significantly improved when using combinations (unigrams + bigrams or mix N-grams).

#### Averaged Results from Three Classification Models: Decision Tree, Naive Bayes and Product Score Based on 15-k Folder Cross Validation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decision Tree</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unigrams</td>
<td>0.570 (0.039)</td>
<td><strong>0.582 (0.071)</strong></td>
<td>0.558 (0.084)</td>
<td>0.570 (0.044)</td>
<td><strong>0.572 (0.044)</strong></td>
<td><strong>0.576 (0.044)</strong></td>
</tr>
<tr>
<td>Bigrams</td>
<td>0.596 (0.045)</td>
<td>0.584 (0.077)</td>
<td>0.607 (0.080)</td>
<td>0.619 (0.069)</td>
<td>0.589 (0.052)</td>
<td></td>
</tr>
<tr>
<td>Trigrams</td>
<td><strong>0.568 (0.033)</strong></td>
<td>0.621 (0.150)</td>
<td>0.512 (0.174)</td>
<td>0.599 (0.053)</td>
<td>0.582 (0.061)</td>
<td></td>
</tr>
<tr>
<td>Uni + Bigrams</td>
<td>0.583 (0.041)</td>
<td>0.598 (0.065)</td>
<td>0.568 (0.085)</td>
<td>0.595 (0.060)</td>
<td>0.564 (0.079)</td>
<td></td>
</tr>
<tr>
<td>Mix N-grams</td>
<td>0.580 (0.037)</td>
<td>0.595 (0.060)</td>
<td>0.564 (0.079)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Naive Bayes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unigrams</td>
<td>0.788 (0.032)</td>
<td>0.779 (0.063)</td>
<td>0.796 (0.071)</td>
<td>0.799 (0.056)</td>
<td>0.801 (0.029)</td>
<td></td>
</tr>
<tr>
<td>Bigrams</td>
<td>0.680 (0.036)</td>
<td>0.888 (0.057)</td>
<td>0.472 (0.110)</td>
<td>0.637 (0.043)</td>
<td>0.832 (0.020)</td>
<td></td>
</tr>
<tr>
<td>Trigrams</td>
<td>0.600 (0.030)</td>
<td><strong>0.922 (0.040)</strong></td>
<td><strong>0.278 (0.078)</strong></td>
<td><strong>0.569 (0.027)</strong></td>
<td>0.827 (0.075)</td>
<td><strong>0.799 (0.018)</strong></td>
</tr>
<tr>
<td>Uni + Bigrams</td>
<td>0.782 (0.039)</td>
<td>0.865 (0.061)</td>
<td>0.699 (0.100)</td>
<td>0.752 (0.056)</td>
<td>0.851 (0.037)</td>
<td><strong>0.799 (0.031)</strong></td>
</tr>
<tr>
<td>Mix N-grams</td>
<td>0.767 (0.032)</td>
<td>0.895 (0.057)</td>
<td>0.640 (0.073)</td>
<td>0.718 (0.036)</td>
<td>0.793 (0.027)</td>
<td></td>
</tr>
<tr>
<td><strong>Product Score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unigrams</td>
<td><strong>0.816 (0.053)</strong></td>
<td>0.852 (0.078)</td>
<td>0.780 (0.073)</td>
<td></td>
<td></td>
<td><strong>0.821 (0.053)</strong></td>
</tr>
<tr>
<td>Bigrams</td>
<td>0.758 (0.043)</td>
<td>0.758 (0.087)</td>
<td>0.769 (0.051)</td>
<td>0.760 (0.053)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trigrams</td>
<td>0.670 (0.050)</td>
<td>0.635 (0.081)</td>
<td>0.781 (0.065)</td>
<td>0.667 (0.073)</td>
<td>0.884 (0.072)</td>
<td></td>
</tr>
<tr>
<td>Uni + Bigrams</td>
<td>0.812 (0.046)</td>
<td>0.814 (0.097)</td>
<td><strong>0.810 (0.054)</strong></td>
<td>0.760 (0.075)</td>
<td>0.810 (0.055)</td>
<td></td>
</tr>
<tr>
<td>Mix N-grams</td>
<td>0.802 (0.051)</td>
<td>0.801 (0.097)</td>
<td>0.804 (0.052)</td>
<td>0.795 (0.059)</td>
<td>0.810 (0.076)</td>
<td>0.800 (0.060)</td>
</tr>
</tbody>
</table>

*Note.* Blue and red represent the lowest and highest value within each column respectively.
Text Classifier Model

- Product Score model using Unigrams
  - Overall prediction accuracy in sample: 81.6%

- Top 10 discriminating keywords:
  - PTSD: emotion, rape, abuse, car, year, flashback, home, nightmare, fire, therapy
  - Non-PTSD: wake, dream, feel, like, anxiety, get, worry, head, breath, sometime
I am 24 years old and was involved in a house fire two years ago. Since then I have split with my long-term boyfriend and not been able to form any other committed relationship. I have been suffering from insomnia regularly, which is impacting on my work situation. I do not think I have flashbacks. I am always aware of where I am but certain smells and sounds make me unable to think about anything else for days at a time and causes me to become really emotional and unable to focus on anything. I have been feeling really disconnected from my life for the last two years and I have finally come to the realization that I need to get help. I am just not sure where to go, or if this is something that will go away on its own.
Example - Output

<table>
<thead>
<tr>
<th>PTSD weight</th>
<th>Non-PTSD weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>('year', 10.076639227930741, 4.8718874052688559)</td>
<td></td>
</tr>
<tr>
<td>('fire', 1.3908600624467784, 0.036088054853843379)</td>
<td></td>
</tr>
<tr>
<td>('year', 10.076639227930741, 4.8718874052688559)</td>
<td></td>
</tr>
<tr>
<td>('flashback', 1.8450184501845019, 0.036088054853843379)</td>
<td></td>
</tr>
<tr>
<td>('emot', 3.8887311950042576, 0.61349693251533743)</td>
<td></td>
</tr>
<tr>
<td>('feel', 14.731762702242406, 25.875135330205701)</td>
<td></td>
</tr>
<tr>
<td>('year', 10.076639227930741, 4.8718874052688559)</td>
<td></td>
</tr>
<tr>
<td>('get', 8.8277036616520004, 14.615662215806568)</td>
<td></td>
</tr>
</tbody>
</table>
Score_PTSDE = 663913.758483
Score_NONPTSD = 17.4703740607
Ratioscore = 10.5454010027

story ( 12 ) with word length: 153 is PTSD and ratioscore is 10.5454010027

===================================================================================================================
Conclusion Study 1: Text mining for PTSD screening

Text mining seems a promising addition to PTSD screening:
- High classification accuracy: high agreement computer vs. clinical diagnoses
- Unigram representation + Product Score model: effective, accurate, balanced sensitivity and specificity

Follow-up research PTSD classification model:
- Apply n-gram analyses on larger, more complex datasets
- Include textual structure features: grammatical properties, part-of-speech
- Extension to Dutch narrative data
- Develop online classification tool
Study 2: Text/audio mining for PTSD treatment

Possibilities:

- Predict development of the PTSD prior to treatment
  - Which treatment fits patients’ needs

- Track patient progress over therapy sessions
  - Change in speaking/writing style
  - Change in structure and organization of the narrative
  - Change in emotions

- Computer-assisted therapy
  - Give feedback to therapists/patients
Current PTSD therapy methods

- Dominant clinical interventions:
  - Cognitive Behavioural Therapy – Brief Eclectic Psychotherapy (BEP)
  - Eye Movement Desensitisation and Reprocessing (EMDR)
  - Psychological debriefing
  - Prolonged exposure/emotional flooding
  - Pharmacotherapy

- E-therapy/Blended therapy:
  - Online or electronic therapy
  - Combination face-to-face and online aid
Brief Eclectic Psychotherapy

- Imaginal exposure to trauma memory by describing trauma
- Focus on “hotspots”: moments with greatest emotional impact

Characteristics of hotspots (Holmes & Grey, 2002)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Recognized by</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Visible) Change in affect</td>
<td>- Burst into tears</td>
</tr>
<tr>
<td></td>
<td>- Turning red</td>
</tr>
<tr>
<td></td>
<td>- Shaking</td>
</tr>
<tr>
<td></td>
<td>- Sweating</td>
</tr>
<tr>
<td>Dissociation</td>
<td>- Change from present to past tense (I am &gt; I was)</td>
</tr>
<tr>
<td></td>
<td>- Change from first to third person (I am &gt; He is)</td>
</tr>
<tr>
<td>Avoidance</td>
<td>- Whizzing through descriptions particular aspects of the trauma</td>
</tr>
<tr>
<td></td>
<td>- Unable to remember details of the moment</td>
</tr>
</tbody>
</table>
Hotspot Classification Framework

Adaption Text Classification Framework He et al., 2013

Phase 1: Training

**LABEL:** Hotspot / Non-hotspot

**Input:**
- Transcripts
- Exposure
- Therapy
- Session

**Feature Extraction Model**

**Features:**
- Linguistic elements
- Phonetic elements
- Structural elements
- Emotions

**Machine Learning Algorithm**

Phase 2: Prediction

**Input:**
- Transcripts
- Exposure
- Therapy
- Session

**Feature Extraction Model**

**Features:**
- Linguistic elements
- Phonetic elements
- Structural elements
- Emotions

**Transcript Classifier Model**

**LABEL:** Hotspot / Non-hotspot
Available dataset

- Audio recordings PTSD patients: N = 45
- BEP therapy – exposure sessions (± 5 p.p.): N = 225
- Coded hotspots for 20 patients (10 worst and 10 best treatments): N = 102
- Recorded time: +/- 45 min per session
- Unique hotspots: N = 64
- Hotspot moments during therapy: N = 95 (31% of the hotspots were repeatedly focused on)

Data processing:
- Literal transcription
- Linking transcription to audio
- Coding variables

UNIVERSITY OF TWENTE.
## Overview variables

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Recognized by</th>
<th>Measure with</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Visible) Change in affect</td>
<td>- Burst into tears</td>
<td>- Emotion detection from Hotspot Manual Holmes &amp; Grey (2002)</td>
</tr>
<tr>
<td></td>
<td>- Turning red</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Shaking</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Sweating</td>
<td></td>
</tr>
<tr>
<td>Dissociation</td>
<td>- Change from present to past tense (I am &gt; I was)</td>
<td>- Content and linguistic analysis</td>
</tr>
<tr>
<td></td>
<td>- Change from first to third person (I am &gt; He is)</td>
<td></td>
</tr>
<tr>
<td>Avoidance</td>
<td>- Whizzing through descriptions particular aspects of the trauma</td>
<td>- Content and linguistic analysis</td>
</tr>
<tr>
<td></td>
<td>- Unable to remember details of the moment</td>
<td>- Phonetic elements: speech rate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Measures for organization/complexity</td>
</tr>
</tbody>
</table>
Hotspot Detection

For available dataset limited to:

- Text analysis:
  - Content
  - Structure

- Audio analysis:
  - Speech rate
  - Pauses between words

- Emotions: coded Hotspots

A tiger and a mouse were walking in a field...
Future ideas

Emotion detection using:
- Phonetic elements
- Facial expressions
Conclusion

- Data Mining promising addition to PTSD screening and treatment
  - Used words in trauma narratives seem to predict risk of PTSD well
  - Track patient progress based on used words
  - Computer assisted therapy (feedback, complete e-therapy)

- Just a first step:
  - More extensive research needed on N-gram analyses needed
  - Validation to other languages necessary
  - Extent model with multimedia (audio/video)

- Aware of difficulties:
  - Text interpretation/speech analysis