Modeling Personal Biases in Language Use by Inducing Personalized Word Embeddings

Daisuke Oba, Naoki Yoshinaga, Shoetsu Sato, Satoshi Akasaki, Masashi Toyoda

June 4, 2019. NAACL-HLT 2019
Word meanings vary depending on each person

What color?
- Yellow
- Golden
- Blonde

How much temperature does the word “Cold” represent?
- ~15°C
- ~20°C
- ~10°C

Personal semantic variations
- Prevent our smooth communication in our daily life
- Degrade the ability for computer to process documents
Goal of our research

Understanding **personal semantic variations**

**RQ1**: Modeling personal sem. var. makes the performance of NLP tasks better?  
* e.g., *sentiment analysis*

**RQ2**: What kind of words have large personal sem. var.?  
* e.g. the words related to the five senses

![Sight | Hearing | Taste | Smell | Touch]

**RQ3**: Common words have low personal sem. var.?
Approach

Compute word embeddings for each person

- e.g., word "Yellow" and "Minneapolis"

![Diagram showing personalized word embeddings]
Existing work

**Goal:** Improve task performances \cite{Tang15, Lin17}

- Obtain personalized word emb. through the task with **subjective output**
  
  e.g., Sentiment analysis

```
Input: Review

Soft mouthfeel and ... finish
High alcohol and ... touch
Drinkable and rich ... good

Estimate

Soft mouthfeel and ... finish
High alcohol and ... touch
Drinkable and rich ... good

Output: Rating
```
Problem on existing work

Bias on output due to subjectivity for each person

😊 Make the task easier to solve without modeling the meanings of input words

-May not learn the meanings

Soft mouthfeel and … finish
High alcohol and … touch
Drinkable and rich … good

Estimate

Mike tend to give higher ratings
Sara tend to give lower ratings

Can not analyze personalized word emb. about their meanings
Idea: Review-target identification task

💡 Task setting that is difficult to solve without modeling meanings of input words (Our target)

Enable personalized word emb. to be analyzed on their meanings

Automatically defined before writing regardless of subjectivity for each person
Overview of our method

**Personalized word embeddings** are jointly trained through the review-target identification task

- Initialized by pretrained Skip-gram embeddings [Mikolov+ 13]

  e.g., review-target: PunkIPA (Beer)

---

![Diagram showing the process of sentence vector smoothing and probabilistic distribution for beer identification.](Image)
**Issue: Data-spariness**

😊 Difficult to stably train the **personalized word emb.** due to the **extremely large number** of review-target
e.g., Amazon review data: \(\sim 10\) million \([\text{He}+, 16]\)
RateBeer dataset: \(\sim 0.1\) million \([\text{McAulley}+, 13]\)
Solution: Multi-task learning (MTL)

Increase supervision by solving relatively easy sub-problems

• Jointly train the metadata identification

<table>
<thead>
<tr>
<th>Review-target</th>
<th>Beer</th>
<th>PunkIPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata</td>
<td>Brewery</td>
<td>Brewdog</td>
</tr>
<tr>
<td></td>
<td>Style</td>
<td>IPA</td>
</tr>
<tr>
<td></td>
<td>ABV</td>
<td>7.0</td>
</tr>
</tbody>
</table>

Solu%on:
Mul%-
LSTM

Beer iden<fica<on

Smooth
mouthfeel
and
finish

Sentence vector

Bi-LSTM

FFNN

Softmax

Probabilistic distribution
of review-target

PunkIPA
Increase supervision by solving relatively easy sub-problems

- Jointly train the metadata identification

<table>
<thead>
<tr>
<th>Review-target</th>
<th>Beer</th>
<th>PunkIPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata</td>
<td>Brewery</td>
<td>Brewdog</td>
</tr>
<tr>
<td></td>
<td>Style</td>
<td>IPA</td>
</tr>
<tr>
<td></td>
<td>ABV</td>
<td>7.0</td>
</tr>
</tbody>
</table>

Solu%on: Mul%‐task learning (MTL)
Two types of experiments

**Intrinsic evaluation**
- **TRAIN**: Review-target identification
- **TEST**: Review-target identification

**Extrinsic evaluation**
- **TRAIN**: review-target identification
- **TEST**: Sentiment analysis
Intrinsic evaluation: settings

Confirm the ability of personalized word emb. to capture meanings of the words for each person

• To successfully solve the task that can’t use output bias, meanings of the input words should be ably modeled

• **Dataset**: RateBeer [McAulley and Leskovec, 13]
  - 526,857 reviews of top-100 users

• **Evaluation**
  - Accuracy (%) of review-target (beer) identification

<table>
<thead>
<tr>
<th></th>
<th>Review-target</th>
<th>Metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beer</td>
<td>(83,049 types)</td>
<td>Brewery (6,208 types)</td>
</tr>
<tr>
<td>Style</td>
<td>(89 types)</td>
<td>ABV</td>
</tr>
</tbody>
</table>
Intrinsic evaluation: result

<table>
<thead>
<tr>
<th>Model</th>
<th>Target identification [Acc.(%)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majority baseline</td>
<td>0.03</td>
</tr>
<tr>
<td>Bi-LSTM</td>
<td>2.99</td>
</tr>
<tr>
<td>+ Multi-task learning with metadata</td>
<td>3.81</td>
</tr>
<tr>
<td>+ Personalization of word embeddings</td>
<td>3.32</td>
</tr>
<tr>
<td>+ ALL</td>
<td>4.14</td>
</tr>
</tbody>
</table>

Capture meanings of input words significantly better than the others (p < 0.01)

Can clarify **personal semantic variations** from the obtained **personalized word embeddings**
Extrinsic evaluation: settings and result

RQ1: Modeling personal sem. var. makes the performance of other NLP tasks better?

• Task: Sentiment analysis with regression
  - Output: Rating scores from 1 to 20

• Evaluation: Root mean square error (RMSE)

<table>
<thead>
<tr>
<th>Model in which the (personalized or not) word embeddings were obtained</th>
<th>Sentiment analysis [RMSE]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bi-LSTM</td>
<td>1.452</td>
</tr>
<tr>
<td>+ Multi-task learning with metadata</td>
<td>1.447</td>
</tr>
<tr>
<td>+ Personalization of word embeddings</td>
<td>1.406</td>
</tr>
<tr>
<td>+ ALL</td>
<td>1.381</td>
</tr>
</tbody>
</table>

👍Personalized word emb. capturing personal sem. var. works better also on an extrinsic task
Analysis: visualization of confused meanings

Two-dimensional representation of the obtained *personalized word embeddings* using PCA

Meanings of the words may be used in different ways by person
Analysis: understanding semantic variation (1)

RQ2: What kind of words have large semantic variations?

Definition of personal semantic variation

- Variance of the personalized word embeddings

| Top -20 | surprisingly, nice, quite, light, pleasant, actually, though, buttery, grassy, really, bready, dusty, fruity, decent, mild, rather, little, toffee, earthy, woody |
| Bottom -20 | lasted, primary, system, secondary, personal, test, acquired, ii, greater, standout, roof, England, flow, scored, purchase, partly, Colorado, spare, rocks, ounce |

The words related to the five senses and adjectives have large semantic var.
Analysis: understanding semantic variation (2)

RQ3: Common words have small semantic variations?

Even if frequently or widely used, meanings of the words are not necessarily agreed among people

• In beer review, frequent words are related to five senses

Pearson correlation 0.55

Pearson correlation 0.51
Summary

**Goal:** Understanding personal semantic variation

**Contribution:** Establish method to eliminate subjective bias on output for each person from the training process of personalized word emb.

**Findings**
- Adjectives and words about the five senses have large sem. var.
- Modeling semantic var. works effectively also on extrinsic task
- Even if the word is common, the meanings are not necessarily agreed

**Future work**
- Further analysis on more datasets and tasks
- Extension of our method for documents other than review text