Query Intent Detection using Convolutional Neural Networks

Homa B. Hashemi, Amir Asiaee, Reiner Kraft

QRUMS workshop - February 22, 2016
Query Intent Detection

michelle obama age → Query Intent Detection → Person Birth Date
Query Intent Detection

michelle obama age

barack obama wife birthday
Roadmap

- Approaches

- Proposed Method
  - Architecture
  - Convolutional Neural Networks model

- Experiments
  - Query Classification
  - Query Clustering
Approaches

- **Rule-Based**
  - Define new rules for a new use case

- **Classification-Based**
  - Feature Engineering
    - Define features
  - Deep Learning Representations
    - Automatically embed queries into a vector space
Deep Learning in NLP

- Deep learning achieved state-of-the-art results in computer vision and speech
- Fast becoming popular in NLP
  - Learning word embeddings (Word2Vec)
  - Learning sentence and document embeddings
    - Semantic parsing (Yih, 2014)
    - Information retrieval (Shen 2014, Huang 2013)
    - Sentiment analysis (RNN and RNTN, Socher et al. 2011, 2012, 2013; Kalchbrenner 2014; Dos Santos 2014; Le 2014; Kim 2014)

- Not much work on Short Text like queries
- Not much work on multi-class classification with high number of classes
Model Architecture

Train Time/Offline

Queries → CNN → Query Vectors → Classifier

Test Time/Online

New Query → CNN → Query Vectors → Classifier → Predicted Intents
Query representations with Word2Vec

1) Get query word vectors
2) Sum or Average vectors
Convolutional Neural Networks

Use pre-trained word vector representations (Google word2vec)

\[ c_i = f(w.x + b) \]
\[ c = [c_1, c_2, ..., c_n] \]
\[ \hat{c} = \max\{c\} \]

Based on Collobert and Weston (2011) and Kim (2014)
# Data

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong># of queries</strong></td>
<td>10,000</td>
</tr>
<tr>
<td><strong># of low-level intent classes</strong></td>
<td>125</td>
</tr>
<tr>
<td><strong># of high-level intent classes</strong></td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>High-level intent</strong></th>
<th><strong>Low-level intent</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Movie</td>
<td>Rating</td>
</tr>
<tr>
<td></td>
<td>Cast</td>
</tr>
<tr>
<td></td>
<td>Length</td>
</tr>
<tr>
<td></td>
<td>Release data</td>
</tr>
<tr>
<td>Person</td>
<td>Spouse</td>
</tr>
<tr>
<td></td>
<td>Birth date</td>
</tr>
<tr>
<td></td>
<td>Children</td>
</tr>
</tbody>
</table>
Query intent detection with 125 low-level intent classes

- 10,000 queries
- 125 intent classes
- 10-fold CV
- Random Forest

Accuracy: 75.06, 81.26, 81.57
Avg Precision: 46, 56, 62
Avg Recall: 39, 45, 47

- Rule-based
- Sum w2v
- Average w2v
- Unigram
- Unigram+bigram
- CNN
Person and Movie entities

- Spouse
- BirthDate
- Height
- Children
- Birth Location
- Movie Cast
- Release Date
- Rating
- Length
- Watch Online

Precision and Recall for:
- Spouse
- Birth Date
- Height
- Children
- Birth Location
- Movie Cast
- Release Date
- Rating
- Length
- Watch Online
Query intent detection with 14 high-level intent classes

- 10,000 queries
- 14 intent classes
- 10-fold CV
- Random Forest

Accuracy
- Rule-based: 77%
- Sum w2v: 89.9%
- Unigram: 90.3%

Precision
- Rule-based: 45
- Sum w2v: 71
- Unigram: 56

Recall
- Rule-based: 34
- Sum w2v: 37
- Unigram: 43
Query Clustering

- K-means
- K = 125
- 10,000 query vector representations

- cast of annie 2014
- big bang cast
- independence day cast and crew
- cast of hollow man
- lee rocker band members
- santana band members
- salif keita band members

- lisa bonet pictures
- bleona qereti photos
- valeen montenegro images
- jennifer aniston wallpapers hd
- alma wade pics
- hitler images

- george selvie wiki
- james brown biography
- bob toski bio

- ninja turtles youtube
- if i stay full movie viooz
- watch homeland online
- the outlander streaming

- john candy death
- ryan knight has died
- is bruce jenner dead
Conclusion

- Use **CNN** to get *query vector representations* for
  - Query Classification
  - Query Clustering

- **Query vector representations**
  - Can be used in all the tasks related to query
  - Can be combined with other methods
    - E.g. combining with rule-based outputs and N-grams to get more discriminative features

- **Future work:**
  - Exploit query vectors using recurrent neural networks with LSTM cells
Detect Emerging intents

- Advantage of CNN over n-grams: Embedding queries into a vector space
- Detect outliers as new set of queries
Detect Emerging intents

- Advantage of CNN over n-grams: Embedding queries into a vector space
- Detect outliers as new set of queries
CNN hyperparameters

- Window filter sizes = 2, 3, 4
- Each filter has 100 feature maps
- Mibi batch size = 50
- Epochs = 40
- Nonlinearity = relu
- Regularization = Dropout
  - Dropout rate = 0.5
  - Randomly set some weights to zero to prevent overfitting