Query Intent Detection using Convolutional Neural Networks

Homa B. Hashemi, Amir Asiaee, Reiner Kraft

QRUMS workshop - February 22, 2016









Jan 04, 2014 · HONOLULU President Barack Obama is going to give his wife the birthday present that many parents can only dream of - time off, alone, without the children.

Roadmap

- Approaches
- Proposed Method
 - Architecture
 - Convolutional Neural Networks model
- Experiments
 - Query Classification
 - > Query Clustering

Approaches



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



Test Time/Online



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



Data

# of queries	10,000
# of low-level intent classes	125
# of high-level intent classes	14

High-level intent	Low-level intent
Movie	Rating
	Cast
	Length
	Release data
Person	Spouse
	Birth date
	Children

Query intent detection with 125 low-level intent classes



Person and Movie entities



Rule-based Unigram CNN



Query intent detection with 14 high-level intent classes

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



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
 - \rightarrow Dropout rate = 0.5
 - > Randomly set some weights to zero to prevent overfitting