YJST at the NTCIR-12 MobileClick-2 Task

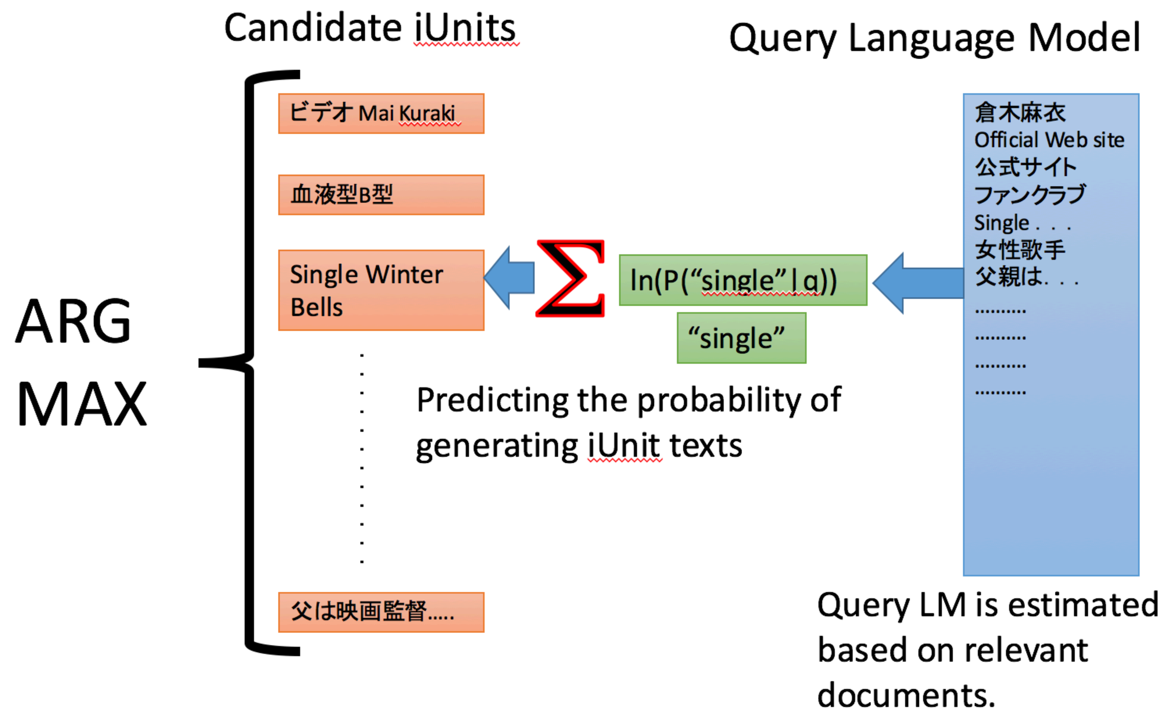
Yuya Ozawa Taichi Yatsuka Sumio Fujita (Yahoo Japan Corp.)
{yozawa,tyatsuka,sufujita}@yahoo-corp.jp

iUnit Ranking

Introduction

We rank each iUnit by a generative model based on Query and background Language Models.

Method



Probability of generating each iUnit is estimated by

- Query LM based on query relevant documents and Background LM based on non relevant documents.
- Models are computed using the title and body fields of provided index data.

Ranking Scores

Dirichlet prior smoothing Uni-gram and Bi-gram models

Uni-gram model

$$P(w|q) = \frac{N_{D_q,w} + \mu P(w|o)}{N_{D_q} + \mu}$$

$$score(u,q) = \sum_{w \in W_u} \ln P(w|q)$$

Bi-gram model

$$P_{bi}(w_{i,i+1}|q) = \begin{cases} \frac{N_{D_q,w_{i,i+1}} + \mu P_{bi}(w_{i,i+1}|o)}{N_{D_q}+\mu} & (w_{i,i+1} \in D_q) \\ \lambda P(w_i|q) & (otherwise) \end{cases}$$

$$score(u,q) = \sum_{w_i, w_{i+1} \in W_u} \ln P_{bi}(w_{i,i+1}|q)$$

Experiments

Training Results

Run description	Run detail	Q-Measure
Random ranking (ORG-R)	–	0.7201
Log Odds Ratio (ORG-L)	Laplace smth	0.7901
Vector Space Cosine	term freq	0.7715
Vector Space Cosine	Boolean	0.78
Vector Space+Background	Boolean	0.8003
Uni-gram Dirichlet priors	$\mu = 1, \alpha = 1$	0.8347
Uni-gram Dirichlet priors	$\mu = 0.5, \alpha = 1$	0.8352
Bi-gram Dirichlet priors	$\mu = 1, \alpha = 0$	**0.8399**
Mixture Dirichlet priors	$\mu = 1, \alpha = 0.5$	0.8375
KL-Divergence	Laplace smth	0.8108
Pitman-Yor	$\mu = 1, \delta = 0.1$	0.8321
iUnit LM	Dir prior $\mu = 1$	0.8258
iUnit LM+cotopic	Dir prior $\mu = 1$	0.8343
iUnit LM+coclick	Dir prior $\mu = 1$	0.8339
iUnit LM+cosession	Dir prior $\mu = 1$	0.8329
iUnit LM+chie	Dir prior $\mu = 1$	0.8345

Test Results

Run description	Run detail	Q-Measure
Random Ranking (ORG-R)	–	0.7411
Log Odds Ratio (ORG-L)	Laplace smth	0.7269
Uni-gram Dirichlet priors	$\mu = 10, \alpha = 1$	**0.8072**
Bi-gram Dirichlet priors	$\mu = 1, \alpha = 0$	0.7965
Mixture Dirichlet priors	$\mu = 1, \alpha = 0.5$	0.8029
Uni-gram Dirichlet priors	$\mu = 0.5, \alpha = 1$	**0.8081**

Conclusions

We used Dirichlet prior smoothing in the LM-Based iUnit ranking approach.

- We carried out several experiments examining : Uni-gram/Bi-gram iUnit/query language models.
- We achieved Q-score of 0.807 in a test run using a Uni-gram model.

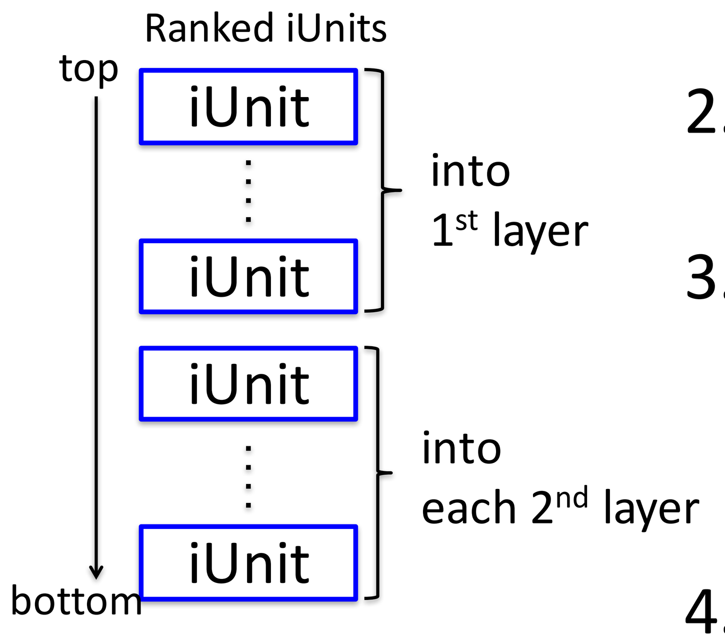
iUnit Summarization

Introduction

We use a word embedding representation to iUnit/intent matching.

Method

LM-Based two layer summarization



1. iUnits are ranked based on the language modeling methods.
2. The top-ranked iUnits are put into the first layer until the length limit X.
3. Lower-ranked iUnits are matched against each intent and put into the second layer of the matched intent according to the **matching score**.
4. Each layer is filled until the text span exceeds the pre-defined length limit.

Matching Scores

Set based Intent Matching (organizer's implementation)

$$Sim_{set}(u,i) = \frac{|W_u \cap W_i|}{|W_i|}$$

**Wx* : the set of words contained in *x*

Word Embedding based Intent Matching (our proposal)

$$Sim_{emb}(u,i) = \cos(Emb_u, Emb_i)$$

$$Emb_u = \sum_{w_u \in W_u} Emb_{w_u}, \quad Emb_i = \sum_{w_i \in W_i} Emb_{w_i}$$

**Embwx* : embedding of the word *wx*

Experiments

Results

Submit #	Run type	Ranking	Intent Matching	Limit	M-measure
123	ORG-T	Log Odds Ratio LM	Set based	280	17.4376
437	Addition	Log Odds Ratio LM	Emb+Cos	280	19.094
131	Official	KL-Div LM	Set based	280	21.0259
173 *1*2	Official	Dir priors LM	Emb+Cos	280	25.8498
231 *2	Official	Dir priors LM	Emb+Cos	0	13.9927
324 *2	Official	Dir priors LM	Emb+Cos	252	25.6084
419 *1	Addition	Dir priors LM	Set based	280	26.7036
442 *1	Addition	Dir priors LM	Emb+Euclidean	280	26.6096

***1** : from #173, #419 & #442

- The vector similarity measure greatly affects the effectiveness of embedding based matching.
- The better usage of word embedding representation hopefully leads to more effective intent matching solutions.

***2** : from #173, #231 & #324

- Reducing the first layer allocation leads to the decline in M-measure.
- It seems that the default limit is near optimum.

Conclusions

- A new intent matching method using word embedding.
  - This leads to a finer allocation of relevant iUnits to subtly related intents in the 2nd layer.
- We achieved M-measure of 25.8498.
  - This is the **best** of official runs of the subtask.
- Further improvement is possible,
  - with more effective similarity matching.