Entity-Duet Neural Ranking: Understanding the Role of Knowledge Graph Semantics in Neural Information Retrieval

Zhenghao Liu, Chenyan Xiong, Maosong Sun, Zhiyuan Liu

Xiaofan Yan

Outline

- Abstract
- Introduction
- Related Work
- Entity-Duet Neural Ranking Model
- Integration with Kernel based Neural Ranking Models
- Experimental Methodology
- Evaluation Results
- Conclusion

Abstract

- EDRM -- Entity-Duet Neural Ranking Model, introduces knowledge graph to neural search systems

Introduction

- Traditional IR
  - entity-oriented search
    - utilizes knowledge graphs to improve search engines
    - Incorporate human knowledge from entities and knowledge semantics
    - Feature-based search system
  - Neural information Retrieval
    - Leverages distributed representations and neural network
    - Sophisticated ranking models from large-scale training data
- Steps
  - Distributed representation: entity/ description/ type embedding
  - Word-entity duet: match documents to queries
  - Feature: no manual but interaction-based neural models

Related work

- NRM(Neural Ranking Model)
  - Representation-based NRM
    - learn good representations and match them in the learned representation space of **query** and **documents**
  - Interaction-based NRM
    - learn word-level interaction patterns from query-document pairs

- Knowledge graph for search system
  - Use KG as pseudo relevance feedback cropus
  - Use KG to build the additional connections

Entity-Duet Neural Ranking Model

- Interaction-based Ranking Model
  - Build translation matrix between q and d, for describe similarities
  - Calculate final ranking score for extract feature on translation matrix

  - first map each word t in q and d to an L-dimensional embedding vector v_t with an embedding layer
  - Then constructs the interaction matrix M, each element in M, compares the i^{th} word in q and the j^{th} word in d using the **cosine similarity** of word embedding:

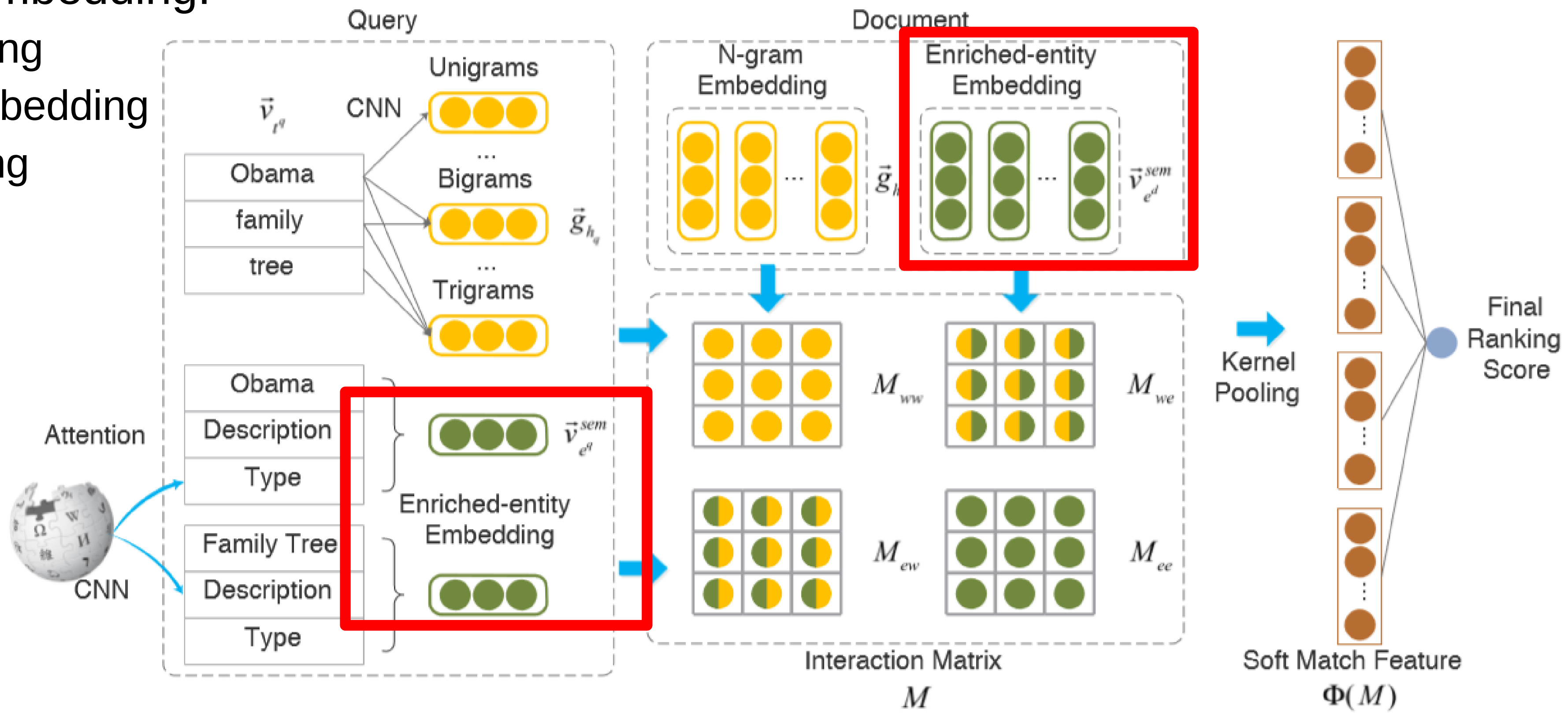
$$M^{ij} = \cos(\vec{v}_{t_i^q}, \vec{v}_{t_j^d}).$$

Entity-Duet Neural Ranking Model

- Semantic Entity Representation
  - KG embedding -- to generate semantic representation
    - Entity embedding: an L^d embedding layer to get entity embedding
    - Description embedding: CNN
    - Type embedding: use an attention mechanism (attention score/ bag-of-words parameter matrix) to combine entity types to type embedding
  - Combination: an linear layer
  - Neural Entity-Duet Framework
    - It utilizes the duet representation of **bag-of-words** and **bag-of-entities** to match q-d with handcrafted features.
    - interaction matrix M: calculates **cosine similarity** between query-document terms
    - Final ranking feature/ final interaction Matrix: concatenation of Ms

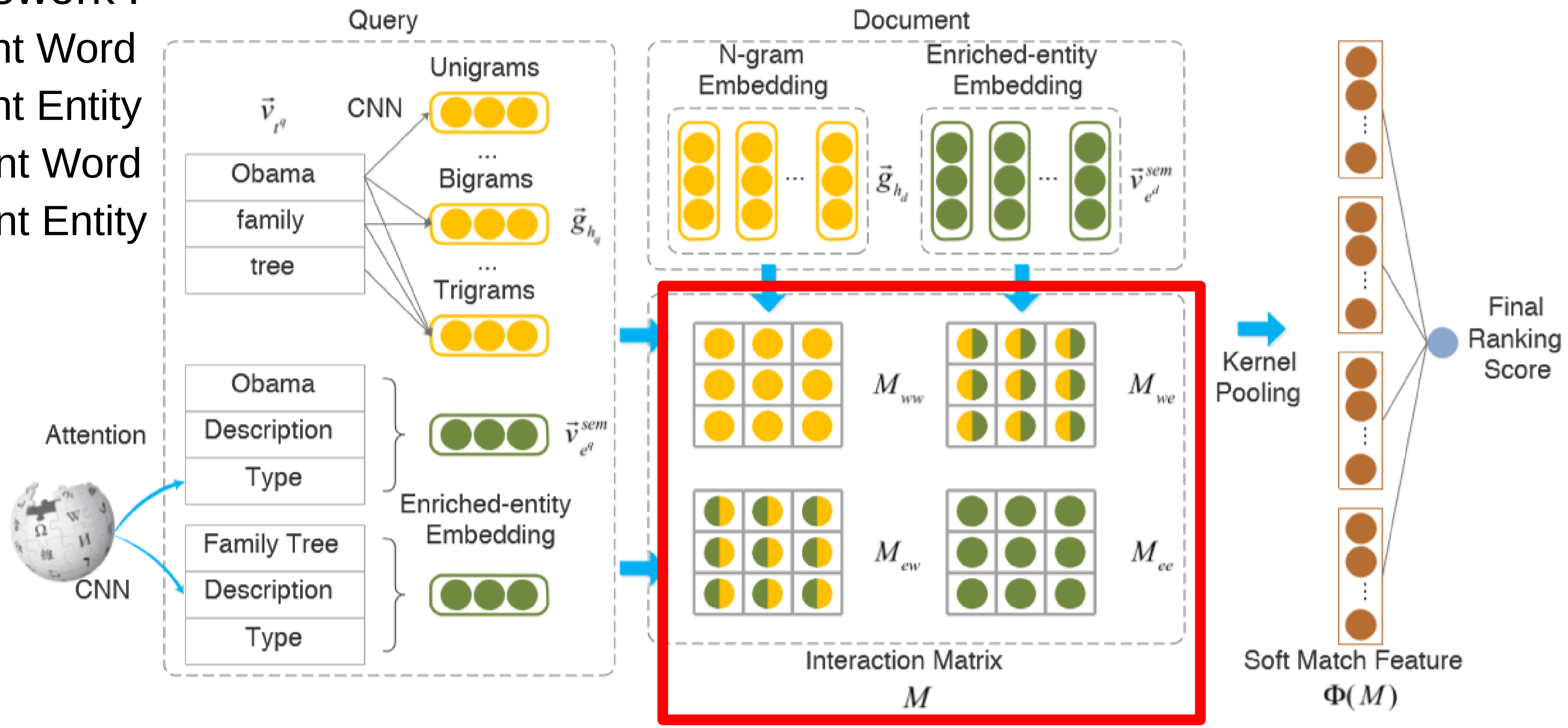
Architecture of EDRM

- Enriched-Entity Embedding:
  - Entity Embedding
  - Description Embedding
  - Type Embedding

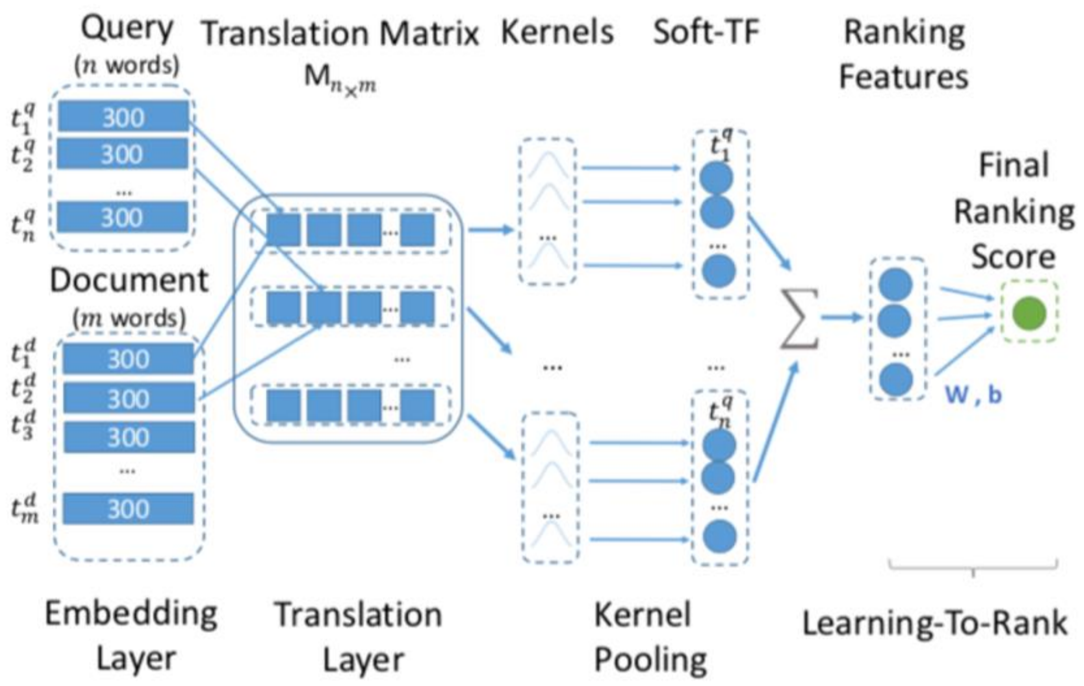


Architecture of EDRM

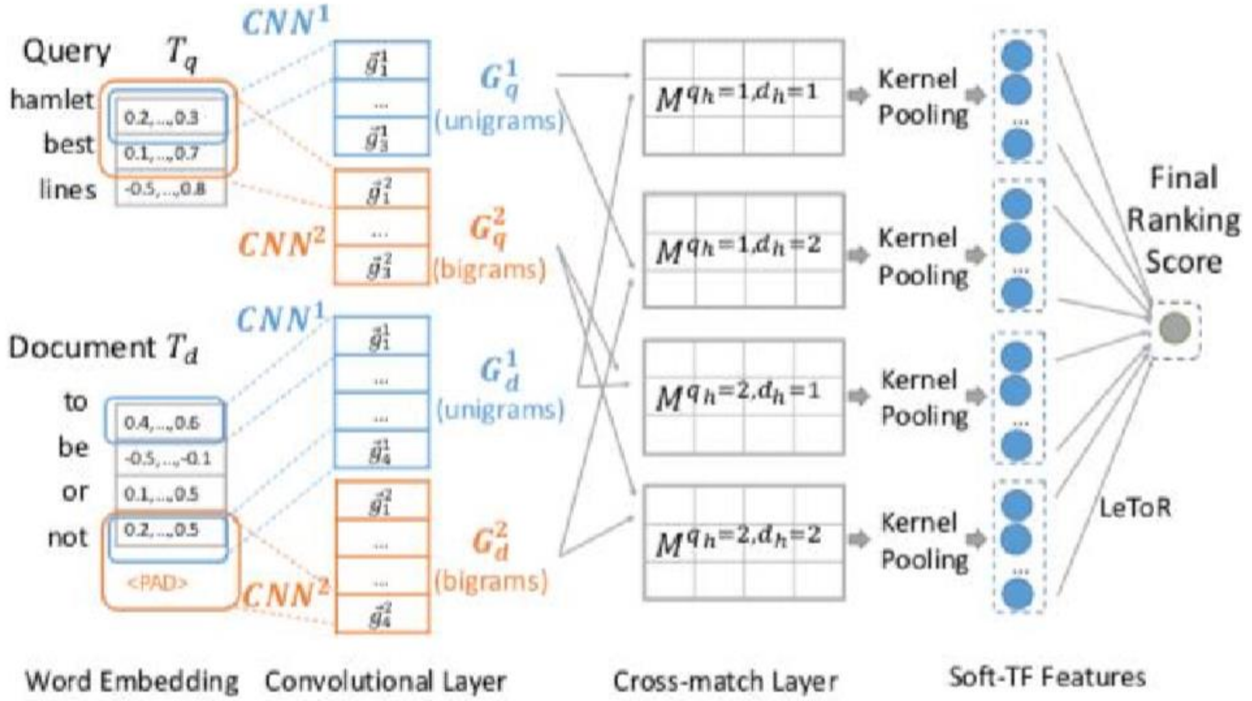
- Neural Entity-Duet Framework :
  - Query Word - Document Word
  - Query Word - Document Entity
  - Query Entity - Document Word
  - Query Entity - Document Entity



Integration with Kernel based Neural Ranking Models



Kernel based Neural Ranking Model
(K-NRM) [Xiong et al., 2017]



Convolutional Kernel-based Neural Ranking Model
(Conv-KNRM) [Dai et al., 2018]

Experimental Methodology

- Dataset: Sougou
  - Click models: model users behaviors to infer relevance signals (labels)
    - DCTR: CTR
    - TACM: both click and dwell time
- Evaluation Metrics

Testing data	Evaluation metrics	
Testing-SAME	NDCG@1	uses DCTR labels
Testing-DIFF	NDCG@10	evaluates models performance based on TACM inferred relevance labels.
Testing-RAW	MRR	evaluates ranking models through user clicks
Statistic significances are tested by permutation test with P< 0.05.		

- Knowledge graph: CN-DBpedia

Experimental Methodology

- Baselines
  - Feature-based ranking models
    - RankSVM (Joachims, 2002)
    - coordinate ascent (Coor-Accent) (Metzler and Croft, 2006)
  - Neural ranking models
    - CDSSM (Shen et al., 2014) -- representation based
    - MatchPyramid (MP) (Pang et al., 2016) -- interaction based
    - DRMM (Grauman and Darrell, 2005) -- interaction based
    - K-NRM (Xiongetal.,2017b)
    - Conv-KNRM (Daietal., 2018)
- Implementation Details

Evaluation Results

- Ranking Accuracy

Table 1: Ranking accuracy of EDRM-KNRM, EDRM-CKNRM and baseline methods. Relative performances compared with K-NRM are in percentages. †, ‡, §, ¶, ∗ indicate statistically significant improvements over DRMM†, CDSSM‡, MP§, K-NRM¶ and Conv-KNRM* respectively.

	Testing-SAME				**Testing-DIFF**				**Testing-RAW**	
Method	**NDCG@1**		**NDCG@10**		**NDCG@1**		**NDCG@10**		**MRR**	
BM25	0.1422	−46.24%	0.2868	−31.67%	0.1631	−45.63%	0.3254	−23.04%	0.2280	−33.86%
RankSVM	0.1457	−44.91%	0.3087	−26.45%	0.1700	−43.33%	0.3519	−16.77%	0.2241	−34.99%
Coor-Ascent	0.1594	−39.74%	0.3547	−15.49%	0.2089	−30.37%	0.3775	−10.71%	0.2415	−29.94%
DRMM	0.1367	−48.34%	0.3134	−25.34%	0.2126‡	−29.14%	0.3592§	−15.05%	0.2335	−32.26%
CDSSM	0.1441	−45.53%	0.3329	−20.69%	0.1834	−38.86%	0.3534	−16.41%	0.2310	−33.00%
MP	0.2184†‡	−17.44%	0.3792†‡	−9.67%	0.1969	−34.37%	0.3450	−18.40%	0.2404	−30.27%
K-NRM	0.2645	–	0.4197	–	0.3000	–	0.4228	–	0.3447	–
Conv-KNRM	0.3357†‡§¶	+26.90%	0.4810†‡§¶	+14.59%	0.3384†‡§¶	+12.81%	0.4318†‡§	+2.14%	0.3582†‡§	+3.91%
EDRM-KNRM	0.3096†‡§¶	+17.04%	0.4547†‡§¶	+8.32%	0.3327†‡§¶	+10.92%	0.4341†‡§¶	+2.68%	0.3616†‡§¶	+4.90%
EDRM-CKNRM	**0.3397**†‡§¶	+28.42%	**0.4821**†‡§¶	+14.86%	**0.3708**†‡§¶*	+23.60%	**0.4513**†‡§¶*	+6.74%	**0.3892**†‡§¶*	+12.90%

Evaluation Results

Kernel? Match kernel?

- Contributions of Matching Kernels
  - Studes the contribution of knowledge graph semantics by investigating the weights learned on the different types of matching kernels.

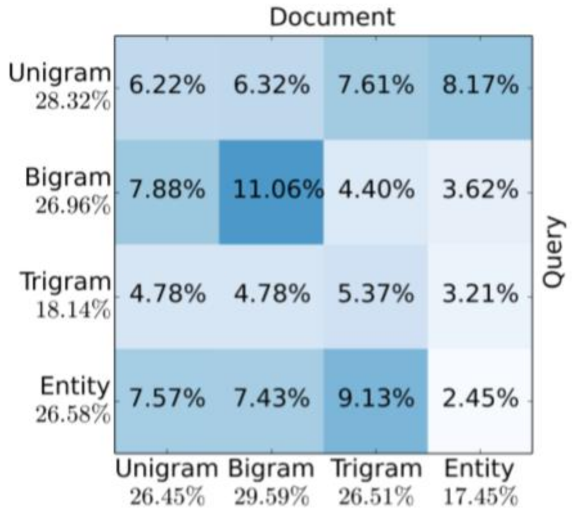
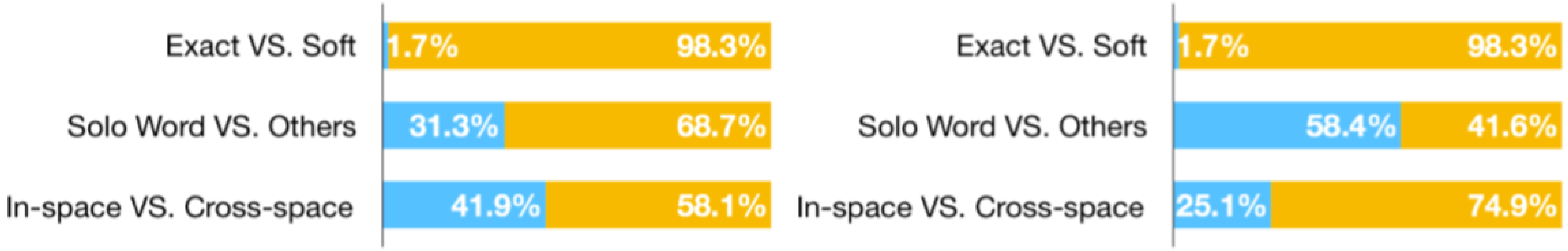


Figure 4: Individual kernel weight for EDRM-CKNRM. X-axis and y-axis denote document and query respectively.



(a) Kernel weight distribution for EDRM-KNRM.

(b) Kernel weight distribution for EDRM-CKNRM.

Figure 3: Ranking contribution for EDRM. Three scenarios are presented: Exact VS. Soft compares the weights of exact match kernel and others; Solo Word VS. Others shows the proportion of only text based matches; In-space VS. Cross-space compares in-space matches and cross-space matches.

Evaluation Results

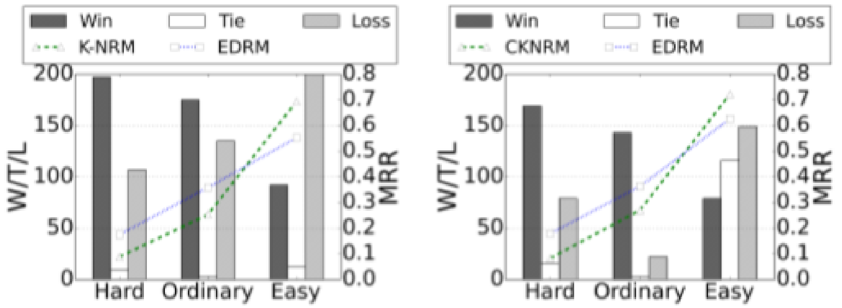
- Ablation Study
  - Studies which part of the knowledge graph semantics leads to the effectiveness and generalization ability of EDRM.
  - **Description** embedding show the greatest improvement.

Table 2: Ranking accuracy of adding diverse semantics based on K-NRM and Conv-KNRM. Relative performances compared are in percentages. †, ‡, §, ¶, ∗, ∗∗ indicate statistically significant improvements over K-NRM† (or Conv-KNRM†), +Embed‡, +Type§, +Description¶, +Embed+Type* and +Embed+Description** respectively.

	Testing-SAME				Testing-DIFF				Testing-RAW	
Method	**NDCG@1**		**NDCG@10**		**NDCG@1**		**NDCG@10**		**MRR**	
K-NRM	0.2645	–	0.4197	–	0.3000	–	0.4228	–	0.3447	–
+Embed	0.2743	+3.68%	0.4296	+2.35%	0.3134	+4.48%	0.4306	+1.86%	0.3641†	+5.62%
+Type	0.2709	+2.41%	0.4395†	+4.71%	0.3126	+4.20%	0.4373†	+3.43%	0.3531	+2.43%
+Description	0.2827	+6.86%	0.4364†	+3.97%	0.3181	+6.04%	0.4306	+1.86%	**0.3691**†§*	+7.06%
+Embed+Type	0.2924†	+10.52%	0.4533†‡§¶	+8.00%	0.3034	+1.13%	0.4297	+1.65%	0.3544	+2.79%
+Embed+Description	0.2891	+9.29%	0.4443†‡	+5.85%	0.3197	+6.57%	0.4304	+1.80%	0.3564	+3.38%
Full Model	**0.3096**†‡§	+17.04%	**0.4547**†‡§¶	+8.32%	**0.3327**†*	+10.92%	**0.4341**†	+2.68%	0.3616†	+4.90%
Conv-KNRM	0.3357	–	0.4810	–	0.3384	–	0.4318	–	0.3582	–
+Embed	0.3382	+0.74%	**0.4831**	+0.44%	0.3450	+1.94%	0.4413	+2.20%	0.3758†	+4.91%
+Type	0.3370	+0.38%	0.4762	−0.99%	0.3422	+1.12%	0.4423†	+2.42%	0.3798†	+6.02%
+Description	0.3396	+1.15%	0.4807	−0.05%	0.3533	+4.41%	0.4468†	+3.47%	0.3819†	+6.61%
+Embed+Type	**0.3420**	+1.88%	0.4828	+0.39%	0.3546	+4.79%	0.4491†	+4.00%	0.3805†	+6.22%
+Embed+Description	0.3382	+0.73%	0.4805	−0.09%	0.3608	+6.60%	0.4494†	+4.08%	0.3868†	+7.99%
Full Model	0.3397	+1.19%	0.4821	+0.24%	**0.3708**†‡§	+9.57%	**0.4513**†‡	+4.51%	**0.3892**†‡	+8.65%

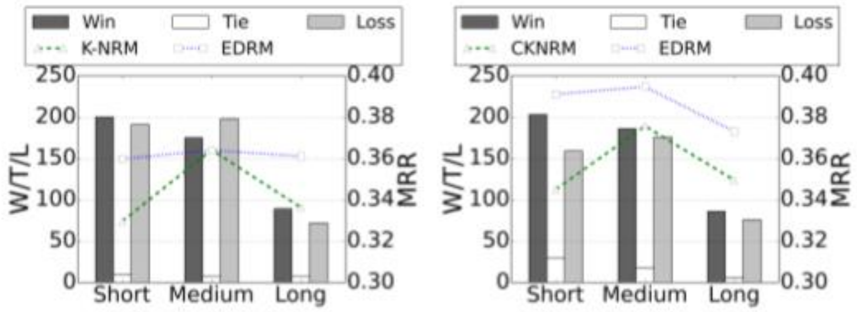
Evaluation Results

- Performance on Different Scenarios
- analyzes the influence of knowledge graphs in two different scenarios: multiple difficulty degrees and multiple length degrees. (搜索的难度和句子长度)



(a) K-NRM VS. EDRM (b) Conv-KNRM VS. EDRM

Figure 5: Performance VS. Query Difficulty. The x-axises mark three query difficulty levels. The y-axises are the Win/Tie/Loss (left) and MRR (right) in the corresponding group.



(a) K-NRM VS. EDRM (b) Conv-KNRM VS. EDRM

Figure 6: Performance VS. Query Length. The x-axises mark three query length levels, and y-axises are the Win/Tie/Loss (left) and MRR (right) in the corresponding group.

- EDRM is more observed on harder or shorter queries

Evaluation Results

- Case Study
  - Knowledge graph semantics could help the document ranking

    1）entity descriptions explain the meaning of entities and connect them through the word space

    2） entity types establish underlying relevance patterns between query and documents.

Table 3: Examples of entity semantics connecting query and title. All the examples are correctly ranked by EDRM-CKNRM. Table 3a shows query-document pairs. Table 3b lists the related entity semantics that include useful information to match the query-document pair. The examples and related semantics are picked by manually examining the ranking changes between different variances of EDRM−CKNRM.

(a) Query and document examples. *Entities* are emphasized.

Query	Document
Meituxiuxiu web version	*Meituxiuxiu web version*: An online picture processing tools
Home page of *Meilishuo*	Home page of *Meilishuo* - Only the correct popular fashion
Master Lu	Master Lu official website: *System optimization*, hardware test, phone evaluation
Crayon Shin-chan: The movie	*Crayon Shin-chan*: The movie online - Anime
GINTAMA	*GINTAMA*: The movie online - Anime - Full HD online watch

(b) Semantics of related entities. The first two rows and last two rows show entity descriptions and entity types respectively.

Entity	Content
Meituxiuxiu web version	Description: Meituxiuxiu is the most popular Chinese image processing software, launched by the Meitu company
Meilishuo	Description: Meilishuo, the largest women's fashion e-commerce platform, dedicates to provide the most popular fashion shopping experience
Crayon Shin-chan, *GINTAMA*	Type: Anime; Cartoon characters; Comic
Master Lu, *System Optimization*	Type: Hardware test; Software; System tool

Evaluation Results

- Contributions of Matching Kernels
  - Using user clicks from search logs, the whole model — the integration of knowledge graph semantics and the neural ranking networks — is trained end-to-end.
  - Leads to a data-driven combination of entity-oriented search and neural information retrieval.
  - Semantics of Knowledge graph -- generalization