

Outline

Morning program

- Preliminaries

- Modeling user behavior

- Semantic matching

- Learning to rank

Afternoon program

- Entities**

- Generating responses

- Recommender systems

- Industry insights

- Q & A

Entities are polysemic

“Finding entities” has multiple meanings.

Entities can be

- ▶ nodes in knowledge graphs,
- ▶ mentions in unstructured texts or queries,
- ▶ retrievable items characterized by texts.

Outline

Morning program

- Preliminaries

- Modeling user behavior

- Semantic matching

- Learning to rank

Afternoon program

Entities

- Knowledge graph embeddings

- Entity mentions in unstructured text

- Entity finding

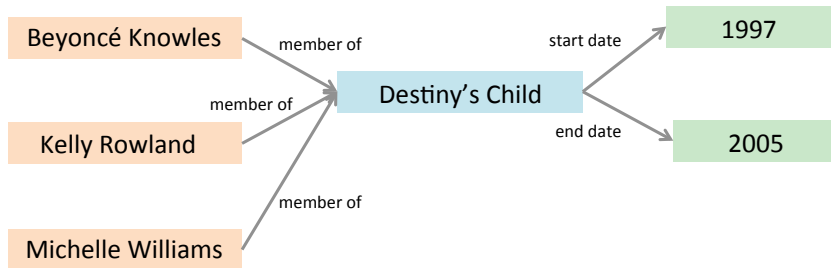
- Generating responses

- Recommender systems

- Industry insights

- Q & A

Knowledge graphs



Triples

```
(beyoncé_knowles, member_of, destinys_child)
(kelly_rowland, member_of, destinys_child)
(michelle_williams, member_of, destinys_child)
(destinys_child, start_date, 1997)
(destinys_child, end_date, 2005)
```

...

Nice overview on using knowledge bases in IR: [Dietz et al., 2017]

Knowledge graphs

Tasks

- ▶ Link prediction

Predict the missing h or t for a triple (h, r, t)

Rank entities by score. Metrics:

- ▶ Mean rank of correct entity
- ▶ Hits@10

- ▶ Triple classification

Predict if (h, r, t) is correct.

Metric: accuracy.

- ▶ Relation fact extraction from free text

Use knowledge base as weak supervision for extracting new triples.

Suppose some IE system gives us $(\text{steve_jobs}, \text{'was the initiator of'}, \text{apple})$, then we want to predict the `founder_of` relation.

Datasets

WordNet

$(\text{car}, \text{hyponym}, \text{vehicle})$

Freebase/DBPedia

$(\text{steve_jobs}, \text{founder_of}, \text{apple})$

Knowlegde graphs

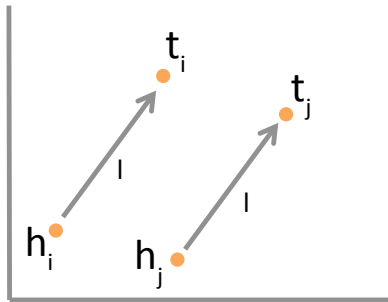
Knowledge graph embeddings

- ▶ TransE [Bordes et al., 2013]
- ▶ TransH [Wang et al., 2014]
- ▶ TransR [Lin et al., 2015]

TransE

“Translation intuition”

For a triple $(h, l, t) : \vec{h} + \vec{l} \approx \vec{t}$.



TransE

“Translation intuition”

For a triple $(h, l, t) : \vec{h} + \vec{l} \approx \vec{t}$.

$$\mathcal{L} = \sum_{(h, l, t) \in S} \sum_{(h', l, t') \in S'_{(h, l, t)}} [\gamma + d(h + l, t) - d(h' + l, t')]_+$$

positive examples negative examples

distance function

[Bordes et al., 2013]

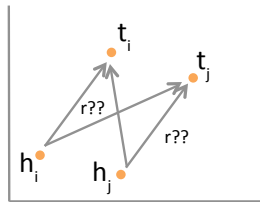
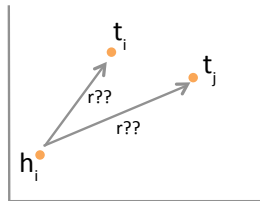
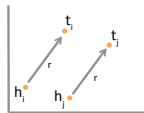
TransE

“Translation intuition”

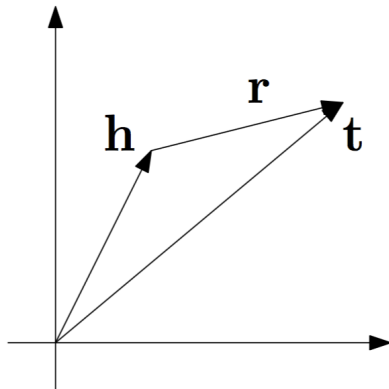
For a triple $(h, l, t) : \vec{h} + \vec{l} \approx \vec{t}$.

How about:

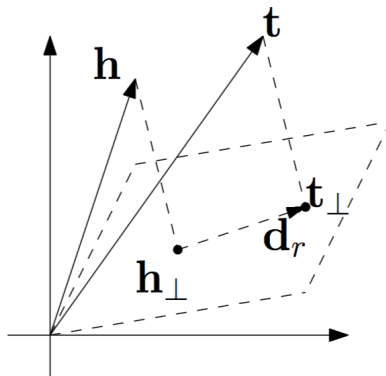
- ▶ one-to-many relations?
- ▶ many-to-many relations?
- ▶ many-to-one relations?



TransH



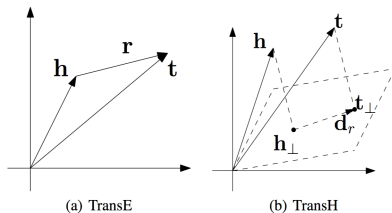
(a) TransE



(b) TransH

[Wang et al., 2014]

TransH



$$f_r(\mathbf{h}, \mathbf{t}) = \|(\mathbf{h} - \mathbf{w}_r^\top \mathbf{h} \mathbf{w}_r) + \mathbf{d}_r - (\mathbf{t} - \mathbf{w}_r^\top \mathbf{t} \mathbf{w}_r)\|_2^2$$

$$\mathcal{L} = \sum_{(h,r,t) \in \Delta} \sum_{(h',r',t') \in \Delta'_{(h,r,t)}} [f_r(\mathbf{h}, \mathbf{t}) + \gamma - f_{r'}(\mathbf{h}', \mathbf{t}')]_+$$

distance function

[Wang et al., 2014]

TransH

Constraints

$$\forall e \in E, \|e\|_2 \leq 1, // \text{scale}$$

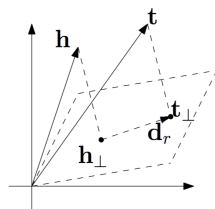
$$\forall r \in R, |\mathbf{w}_r^\top \mathbf{d}_r| / \|\mathbf{d}_r\|_2 \leq \epsilon, // \text{orthogonal}$$

$$\forall r \in R, \|\mathbf{w}_r\|_2 = 1, // \text{unit normal vector}$$

i.e., translation vector \mathbf{d}_r
is in the hyperplane

$$\mathcal{L} = \sum_{(h,r,t) \in \Delta} \sum_{(h',r',t') \in \Delta'_{(h,r,t)}} [f_r(\mathbf{h}, \mathbf{t}) + \gamma - f_{r'}(\mathbf{h}', \mathbf{t}')]_+ \quad \text{soft constraints}$$

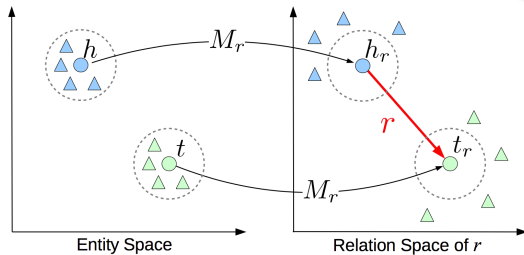
$$+ C \left\{ \sum_{e \in E} [\|e\|_2^2 - 1]_+ + \sum_{r \in R} \left[\frac{(\mathbf{w}_r^\top \mathbf{d}_r)^2}{\|\mathbf{d}_r\|_2^2} - \epsilon^2 \right]_+ \right\}, \quad (4)$$



(b) TransH

[Wang et al., 2014]

TransR

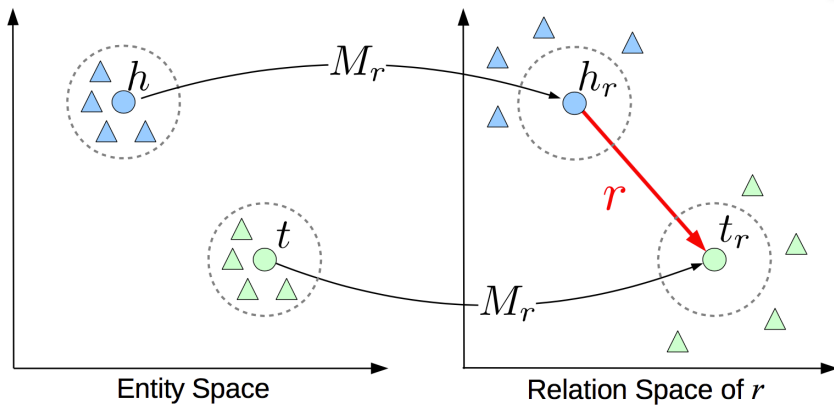


Use different embedding spaces for entities and relations

- ▶ 1 entity space
- ▶ multiple relation spaces
- ▶ perform translation in appropriate relation space

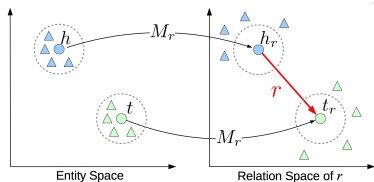
[Lin et al., 2015]

TransR



[Lin et al., 2015]

TransR



Relations: R^d

Entities: R^k

M_r = projection matrix: $k * d$

$$\mathbf{h}_r = \mathbf{h}M_r, \quad \mathbf{t}_r = \mathbf{t}M_r$$

$$f_r(h, t) = \|\mathbf{h}_r + \mathbf{r} - \mathbf{t}_r\|_2^2$$

Constraints:

$$\|\mathbf{h}\|_2 \leq 1$$

$$\|\mathbf{r}\|_2 \leq 1$$

$$\|\mathbf{t}\|_2 \leq 1$$

$$\|\mathbf{t}M_r\|_2 \leq 1$$

$$\|\mathbf{h}M_r\|_2 \leq 1$$

$$L = \sum_{(h,r,t) \in S} \sum_{(h',r,t') \in S'} \max(0, f_r(h, t) + \gamma - f_r(h', t'))$$

[Lin et al., 2015]

Challenges

- ▶ How about time?
E.g., some relations hold from a certain date, until a certain date.
- ▶ New entities/relationships
- ▶ Finding synonymous relationships/duplicate entities (2005, end_date, destinys_child) (destinys_child, disband, 2005) (destinys_child, last_performance, 2005)
- ▶ Evaluation
Link prediction? Relation classification? Is this fair? As in, is this even possible in all cases (for a human without any world knowledge)?

Resources: toolkits + knowledge bases

Source Code

KB2E : <https://github.com/thunlp/KB2E> [Lin et al., 2015]

TransE : <https://everest.hds.utc.fr/>

Knowledge Graphs

- ▶ Google Knowledge Graph
google.com/insidesearch/features/search/knowledge.html
- ▶ Freebase
freebase.com
- ▶ GeneOntology
geneontology.org
- ▶ WikiLinks
code.google.com/p/wiki-links

Outline

Morning program

- Preliminaries

- Modeling user behavior

- Semantic matching

- Learning to rank

Afternoon program

Entities

- Knowledge graph embeddings

- Entity mentions in unstructured text

- Entity finding

- Generating responses

- Recommender systems

- Industry insights

- Q & A

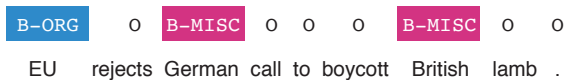
Entity mentions

Recognition Detect mentions within unstructured text (e.g., query).

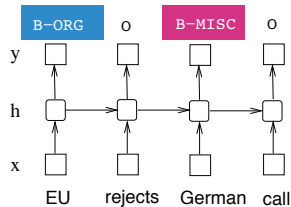
Linking Link mentions to knowledge graph entities.

Utilization Use mentions to improve search.

Named entity recognition



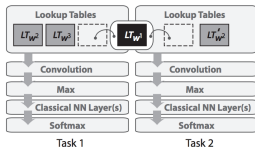
Task



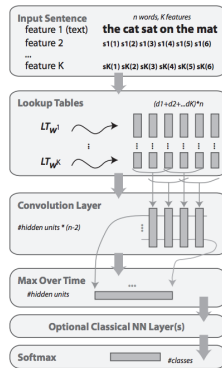
vanilla RNN

Named entity recognition

- ▶ A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning [Collobert and Weston, 2008]
- ▶ Natural Language Processing (Almost) from Scratch [Collobert et al., 2011]

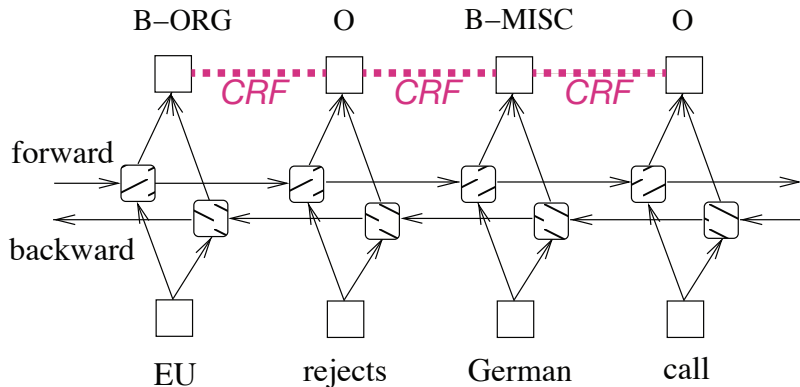


Learning a single model to solve multiple NLP tasks. Taken from [Collobert and Weston, 2008].



Feed-forward language model architecture for different NLP tasks. Taken from [Collobert and Weston, 2008].

Named entity recognition

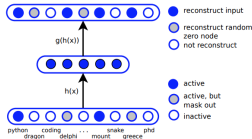


BI-LSTM-CRF model

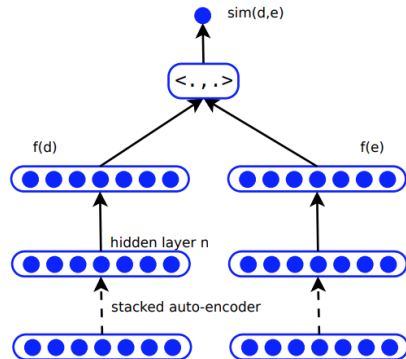
[Huang et al., 2015]

Entity disambiguation

- ▶ Learn representations for documents and entities.
- ▶ Optimize a distribution of candidate entities given a document using (a) cross entropy or (b) pairwise loss.

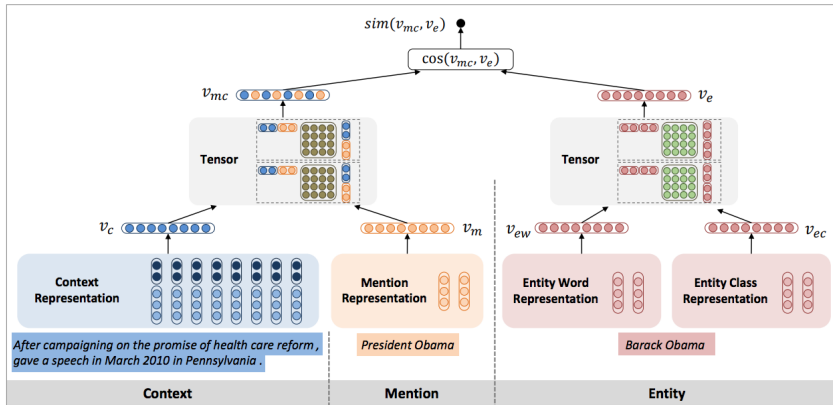


Learn initial document representation in unsupervised pre-training stage. Taken from [He et al., 2013].



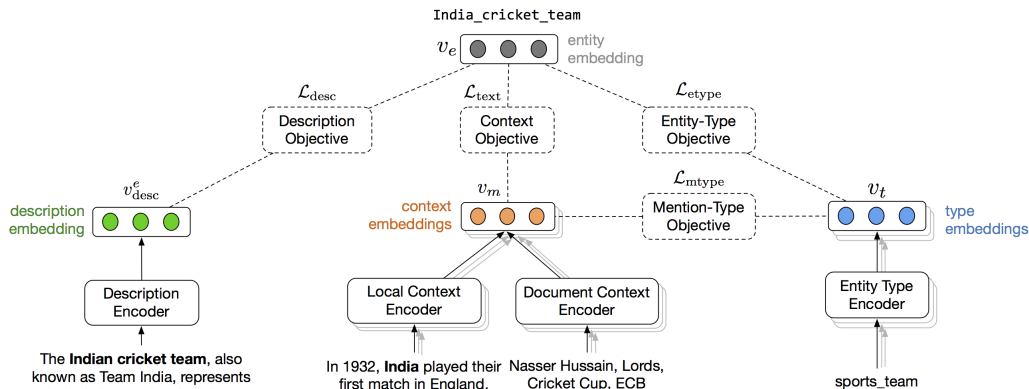
Learn similarity between document and entity representations using supervision. Taken from [He et al., 2013].

Entity linking



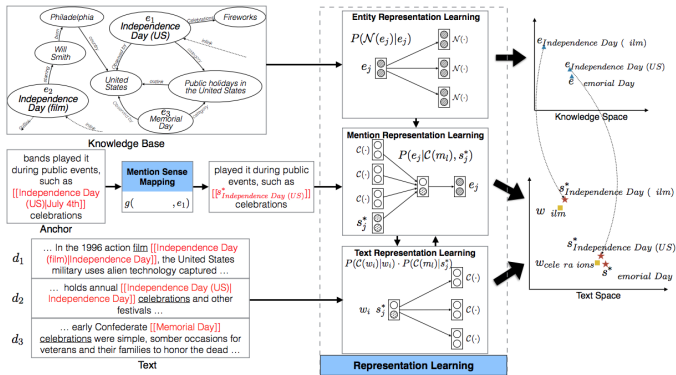
Learn representations for the context, the mention, the entity (using surface words) and the entity class. Uses pre-trained word2vec embeddings. Taken from [Sun et al., 2015].

Entity linking



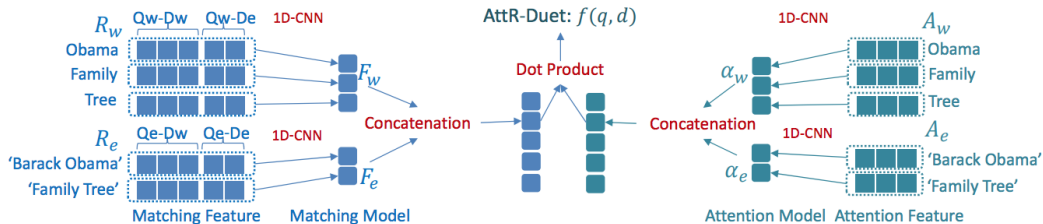
Encode Wikipedia descriptions, linked mentions in Wikipedia and fine-grained entity types. All representations are optimized jointly. Taken from [Gupta et al., 2017].

Entity linking



A single mention phrase refers to various entities. Multi-Prototype Mention Embedding model that learns multiple sense embeddings for each mention by jointly modeling words from textual contexts and entities derived from a KB. Taken from [Cao et al., 2017].

Improving search using linked entities



Attention-based ranking model for word-entity duet. Learn a similarity between query and document entities. Resulting model can be used to obtain ranking signal. Taken from [Xiong et al., 2017a].

Outline

Morning program

- Preliminaries

- Modeling user behavior

- Semantic matching

- Learning to rank

Afternoon program

Entities

- Knowledge graph embeddings

- Entity mentions in unstructured text

Entity finding

- Generating responses

- Recommender systems

- Industry insights

- Q & A

Entity finding

Task definition

Rank entities satisfying a topic described by a few query terms.

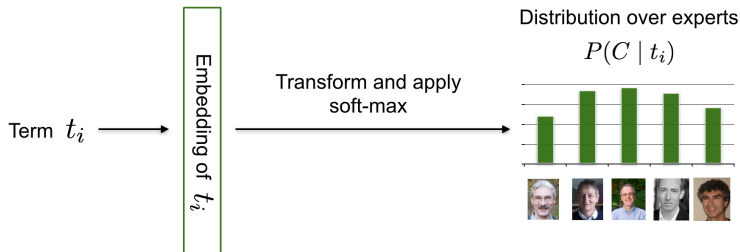
Not just document search — (a) topics do not typically correspond to entity names, (b) average textual description much longer than typical document.

Different instantiations of the task within varying domains:

- ▶ Wikipedia: INEX Entity Ranking Track [de Vries et al., 2007, Demartini et al., 2008, 2009, 2010] (lots of text, knowledge graph, revisions)
- ▶ Enterprise search: expert finding [Balog et al., 2006, 2012] (few entities, abundance of text per entity)
- ▶ E-commerce: product ranking [Rowley, 2000] (noisy text, customer preferences)

Semantic Expertise Retrieval [Van Gysel et al., 2016]

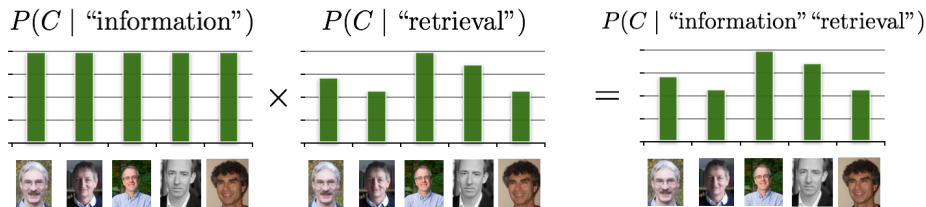
- ▶ Expert finding is a particular entity retrieval task where there is a lot of text.
- ▶ Learn representations of words and entities such that n-grams extracted from a document predict the correct expert.



Taken from slides of Van Gysel et al. [2016].

Semantic Expertise Retrieval [Van Gysel et al., 2016] (cont'd)

- ▶ Expert finding is a particular entity retrieval task where there is **a lot** of text.
- ▶ Learn representations of words and entities such that n-grams extracted from a document predict the correct expert.



Taken from slides of Van Gysel et al. [2016].

Regularities in Text-based Entity Vector Spaces [Van Gysel et al., 2017c]

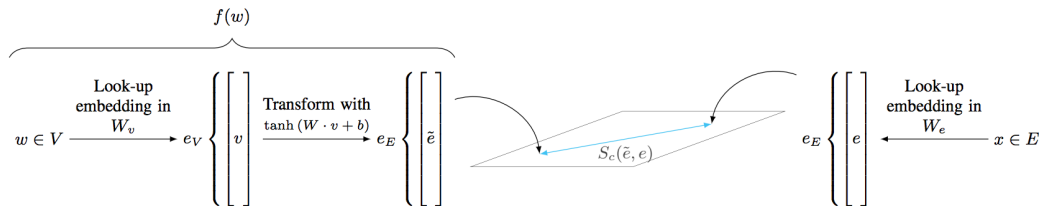
To what extent do entity representation models, trained only on text, encode structural regularities of the entity's domain?

Goal: give insight into learned entity representations.

- ▶ Clusterings of experts correlate somewhat with groups that exist in the real world.
- ▶ Some representation methods encode co-authorship information into their vector space.
- ▶ Rank within organizations is learned (e.g., Professors $>$ PhD students) as senior people typically have more published works.

Latent Semantic Entities [Van Gysel et al., 2016]

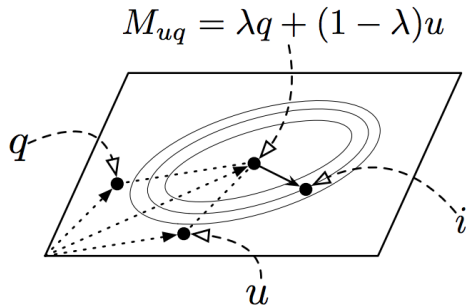
- ▶ Learn representations of e-commerce products and query terms for product search.
- ▶ Tackles learning objective scalability limitations from previous work.
- ▶ Useful as a semantic feature within a Learning To Rank model in addition to a lexical matching signal.



Taken from slides of Van Gysel et al. [2016].

Personalized Product Search [Ai et al., 2017]

- ▶ Learn representations of e-commerce products, query terms, and users for personalized e-commerce search.
- ▶ Mixes **supervised** (relevance triples of query, user and product) and **unsupervised** (language modeling) objectives.
- ▶ The query is represented as an interpolation of query term and user representations.



Personalized product search in a latent space with query \vec{q} , user \vec{u} and product item \vec{i} . Taken from Ai et al. [2017].

Resources: toolkits

SERT : <http://www.github.com/cvangysel/SERT> [Van Gysel et al., 2017b]

HEM : <https://ciir.cs.umass.edu/downloads/HEM> [Ai et al., 2017]

Resources: further reading on entities/KGs

For more information, see the tutorial on “Utilizing Knowledge Graphs in Text-centric Information Retrieval” [Dietz et al., 2017] presented at last year’s WSDM.

<https://github.com/laura-dietz/tutorial-utilizing-kg>