OxKBC
Outcome Explanation for Factorization Based Knowledge Base Completion

Yatin Nandwani, Ankesh Gupta, Aman Agrawal, Mayank S. Chauhan, Parag Singla & Mausam
Department of Computer Science & Engineering
Indian Institute of Technology Delhi, INDIA
Many accurate TF models for KBC, e.g. DistMult*, ComplEx#
Many accurate TF models for KBC, e.g. DistMult*, ComplEx#

Models are opaque
  - Reduces user trust in prediction

Contribution

- Many accurate TF models for KBC, *e.g.* DistMult*, ComplEx#*
- Models are opaque
  - Reduces user trust in prediction
- **OxKBC**: Generate explanations for model predictions

#Trouillon Théo et al. "Complex embeddings for simple link prediction." ICML, 2016."
**Contribution**

- Many accurate TF models for KBC, *e.g.* DistMult*, ComplEx#*
- Models are opaque
  - Reduces user trust in prediction
- **OxKBC**: Generate explanations for model predictions
  - Trains a *post-facto* explanator

---

#Trouillon Théo et al. "Complex embeddings for simple link prediction." ICML, 2016.*
**Contribution**

- Many accurate TF models for KBC, *e.g.* DistMult*, ComplEx#*
- Models are opaque
  - Reduces user trust in prediction
- **OxKBC**: Generate explanations for model predictions
  - Trains a *post-facto* explanator
  - Faithful to the model

**Contribution**

- Many accurate TF models for KBC, *e.g.* DistMult*, ComplEx#*
- Models are opaque
  - Reduces user trust in prediction
- **OxKBC**: Generate explanations for model predictions
  - Trains a *post-facto* explainer
  - Faithful to the model
  - Satisfies end user in a user study on MTurk

#Trouillon Théo et al. "Complex embeddings for simple link prediction." ICML, 2016.*
What is a TF Model?

- Knowledge Graph:
  - Nodes: Entities e.g. Tory, BasketBall
  - Directed Edges: Relations b/w edges e.g. GoodAt
What is a TF Model?

- **Knowledge Graph:**
  - Nodes: Entities *e.g.* Tory, BasketBall
  - Directed Edges: Relations b/w edges *e.g.* GoodAt

- **TF Model:** Learn embeddings for each entity and relation
What is a TF Model?

- **Knowledge Graph:**
  - **Nodes:** Entities e.g. Tory, BasketBall
  - **Directed Edges:** Relations b/w edges e.g. GoodAt

- **TF Model:** Learn embeddings for each entity and relation
OxKBC: Outcome explanation for KBC

Query

< $s, r, ? >$

Prediction

< $s, r, o >$

TF Model

GoodAt

Basketball
OxKBC: Outcome explanation for KBC

Query

< s, r, ? >

TF Model

 prediction

< s, r, o >

OxKBC

GoodAt

Basketball
OxKBC: Outcome explanation for KBC

Query

< s, r, ? >

TF Model

Prediction

< s, r, o >

OxKBC

Explanation

GoodAt

Query

< s, r, ? >

TF Model

Prediction

< s, r, o >

OxKBC

Explanation

GoodAt
OxKBC: Outcome explanation for KBC

Query

< s, r, ? >

TF Model

Prediction

< s, r, o >

OxKBC

Explanation

< s, r, ? >

because

< s, r', ? > & r' ~ r
Taxonomy of Explanation Engines

Source credits: On explainable AI: From theory to motivation, applications and limitations, Tutorial at AAAI 2019
OxKBC: Different Templates of Explanations
OxKBC: Different Templates of Explanations

- Entity Similarity (T1)
OxKBC: Different Templates of Explanations

- Entity Similarity (T1)
- Relation Similarity (T2)
OxKBC: Different Templates of Explanations

- Entity Similarity (T1)
- Relation Similarity (T2)
- Two Length Paths (T3)
OxKBC: Selection Module

- How to pick the best explanation?
OxKBC: Selection Module

- How to pick the best explanation?
- No annotation over explanations.
OxKBC: Selection Module

- How to pick the best explanation?
- No annotation over explanations.
- Selection Module:
OxKBC: Selection Module

- How to pick the best explanation?
- No annotation over explanations.
- **Selection Module:**
  - Novel features for each template
OxKBC: Selection Module

- How to pick the best explanation?
- No annotation over explanations.
- **Selection Module:**
  - Novel features for each template
  - Novel unsupervised loss function
OxKBC: Selection Module

- How to pick the best explanation?
- No annotation over explanations.
- **Selection Module:**
  - Novel features for each template
  - Novel unsupervised loss function
  - Handful of annotation reduces variance
Turk Experiments

○ How good are OxKBC’s explanations?
○ Do end users prefer OxKBC vs Rule Mining*?

Turk Experiments

Question: Turkun Palloseura football team has a play position ______?

Answer: Midfielder

Explanation A: Turkun Palloseura football team has a play position Defender and Defender is a position in the football team Lierse S.K. and Lierse S.K. has a position of Midfielder

Explanation B: In our Knowledge Base, many football teams (598 of 745) have players at Midfielder

(Australia national soccer team, Olympique de Marseille and 596 more...) football team has a play position Midfielder

- A is better than B
- B is better than A
- Both A and B are equally good
- Both A and B are bad
## Turk Experiments

<table>
<thead>
<tr>
<th></th>
<th>OxKBC Better</th>
<th>Rules Better</th>
<th>Tie</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>145</strong></td>
<td>49</td>
<td>18</td>
<td><strong>212</strong></td>
</tr>
</tbody>
</table>
Faithfulness of Explanations

- What if explaining facts are removed from KB?
Faithfulness of Explanations

- What if explaining facts are removed from KB?
- How will it impact model’s prediction?
Faithfulness of Explanations

- What if explaining facts are removed from KB?
- How will it impact model’s prediction?
- Measure through reduction in MRR:
  - More decrease => More faithful
Faithfulness of Explanations

- What if explaining facts are removed from KB?
- How will it impact model’s prediction?
- Measure through reduction in MRR:
  - More decrease => More faithful

<table>
<thead>
<tr>
<th>Steps</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>0.50</td>
<td>0.60</td>
<td>0.40</td>
</tr>
<tr>
<td>2</td>
<td>0.30</td>
<td>0.34</td>
<td>0.24</td>
</tr>
<tr>
<td>5</td>
<td>0.20</td>
<td>0.18</td>
<td>0.13</td>
</tr>
</tbody>
</table>
Conclusion

- OxKBC provides post-hoc explanations
  - for any factorization based KBC models.
- Faithful to the underlying model
- Satisfies end user in a user study on MTurk
- Increases trust in the underlying model

All code and data at: https://github.com/dair-iitd/OxKBC