On embeddings as an alternative relational learning paradigm

Sebastijan Dumančić, Alberto García-Durán, Mathias Niepert
sebastian.dumancic@cs.kuleuven.be, {alberto.duran.mathias.niepert}@neclab.eu
KU Leuven, Belgium; NEC Labs Europe, Germany

Paradigms of relational learning

Statistical relational learning
(Probabilistic) inductive logic programming
→ use predicate logic (and probability theory) to represent complex data

Knowledge graph embeddings
→ re-represent data as vectors

What are the relative strengths of the two paradigms?

Goal: compare the typical representatives of ILP (because of the better support for learning the logical theories) and KG embeddings on a series of classification and clustering tasks; also include relational latent representation learning approach CUR²LED

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Reasoning depth</th>
<th>Rule properties</th>
<th>Dataset properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hepatitis</td>
<td>1.4</td>
<td>10 %</td>
<td>36 %</td>
</tr>
<tr>
<td>Terrorsite</td>
<td>1</td>
<td>0 %</td>
<td>96 %</td>
</tr>
<tr>
<td>Mutagenesis</td>
<td>0</td>
<td>0 %</td>
<td>100 %</td>
</tr>
<tr>
<td>WebKB</td>
<td>2</td>
<td>5 %</td>
<td>76 %</td>
</tr>
</tbody>
</table>

Hyper-parameters of the embeddings methods are very important

Performance is very sensitive to the choice of the dimension of the embedding vectors

Unfortunately, the clustering performance is extremely sensitive to the dimension of the embedding vectors

How to make this choice without the labels? No clear way to do that