

Improving Entity Linking by Modeling Latent Relations between Mentions

Phong Le, Ivan Titov

Xiaofan Yan

Innovation

- Relations
 - exploit relations between textual mentions in a document (e.g., coreference) to decide if the linking decisions are compatible
 - Previous model:
 - relied on supervised systems/heuristics to predict relations
 - this model:
 - treat relations as latent variables

Background

- Local model:
 - Local models rely only on local contexts of mentions and completely ignore interdependencies between the linking decisions in the document. $\Psi(ei, ci)$ as a **local score** function
- Global model:
 - A global model, besides using local context within $\Psi(ei, ci)$, takes into account entity coherency. It is captured by a **coherence score** function $\Phi(E, D)$:

Model

- Relation:
 - relation k is assigned to a mention pair (m_i, m_j) with a non-negative weight ('confidence') α_{ijk} . The pairwise score (m_i, m_j) is computed as a weighted sum of relation-specific pairwise scores
- Pairwise score:
 - Pairwise score function: Ganea and Hofmann (2017)
 - Normalization:
 - Rel-norm (Relation-wise normalization): coefficients α_{ijk} are normalized over relations, and use a small fixed number of relations as relation embedding
 - Ment-norm (Mention-wise normalization): multi-head attention mechanism
- Mention padding:
 - add to each document a padding mention linked to a padding entity

Model

- Implementation
 - Define a CRF, using a local score function to define pairwise score
 - Use max-product LBP to estimate the max-marginal probability
 - Generate final score by A two-layer neural network
 - Loss