|  |  |
| --- | --- |
| Title | Distant supervision for relation extraction without labeled data |
| Author | Mike Mintz, Steven Bills, Rion Snow, Dan Jurafsky |
| URL | https://www.aclweb.org/anthology/P09-1113 |
| Content | 3 kinds of learning paradigms have been applied to the task of extracting relational facts from text:  - supervised approaches  - purely unsupervised information extraction  - boot-strap learning  + our new method: distant supervision  We investigate an alternative paradigm that does not require labeled  corpora, avoiding the domain dependence of ACE(Automatic Content Extraction) style algorithms, and allowing the use of corpora of any size.  One of the main advantages of our architecture is its ability to combine information from many different mentions of the same relation.  Supervised IE: supervised information extraction  The **distant supervision assumption** is that if two entities participate in a relation, any sentence that contain those two entities might express that relation.  **Architecture:**  In the training step, all entities are identified 1005 in sentences using a named entity tagger that labels persons, organizations and locations. If a sentence contains two entities and those entities are an instance of one of our Freebase relations, features are extracted from that sentence and are added to the feature vector for the relation.  **Evaluation:**  automatically and manually |
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|