|  |  |
| --- | --- |
| Title | CERES: Distantly Supervised Relation Extraction from the Semi-Structured Web |
| Author | Colin Lockard, Xin Luna Dong, Arash Einolghozati, Prashant Shiralkar |
| URL | http://www.vldb.org/pvldb/vol11/p1084-lockard.pdf |
| Content | Distant supervision does not easily apply to semi-structured websites because the unit for annotation changes from a sentence to a webpage, normally containing much more information.**Too much information**  Contributions:  - describe CERES, an end-to-end distantly supervised knowledge extraction framework that can apply to semi-structured websites independent of domains and web sources.  - We propose an advanced distant supervision annotation process for semi-structured web sources by leveraging the unique structural characteristics of such data.  - In addition to showing the efficacy of our technique on a benchmark data set, we include an evaluation on dozens of real-world long tail websites containing multi-valued predicates to show applicability of our method in the wild. |
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|
|