|  |  |  |  |
| --- | --- | --- | --- |
| **Paper Serial** | R5 – S – NER | | |
| **Paper Name** | ZeroShotCeres: Zero-Shot Relation Extraction from Semi-Structured Webpages | | |
| **Conference/Journal Name** | arXiv.org | | |
| **Journal Category** | W | **Year of Publication** | 2021 |
| **Language** | English |
| **Objective of the Research** | Introduce a method for Relation Extraction from webpages | | |
| **Major Contribution/ Gaps Addressed** | Create a task of relation extraction (named as ZeroShot) and a GNN model to extract information from documents which never-before-seen.  The method works in both OpenIE and ClosedIE settings. | | |
| **Approach/ Method/ Technique** | Propose the first approach to enable Open Information Extraction from semi-structured websites without prior knowledge or training data in the subject. A Deep Learning technique (GNN) was used for webpage represenation that integrates multi-modal information including visual, layout, and textual features enabling generalize IE from never-before-seen websites. | | |
| **Application Domain** | Open Domain | | |
| **Data Set Details** | For both OpenIE & CloseIE, SWDE dataset was used.  Which contains labels for OpenIE extractions of 21 English Language websites (each with one template) in threee subjects (Movie, NBA and University) with between 400 and 2000 pages per site.  ClosedIE labels generated by converting OpenIE labels via manual alignment of OpenIE relations between websites, giving a set of 18 relations for Movie vertical, 14 for NBA and 13 for University. | | |
| **Experimental Setup** | The training setting classify into following categories, indicating the level of vertical or site-specific knowledge used, in decreasing level of difficulty.  Level 1: Unseen-Vertical Zero-shot, it's build only for OpenIE, a model is trained on sites from two of the three verticals and applied to sites from other vertical.  Level 2: Zero-shot with Vertical Knowledge, a model is trained on all sites but one (spanning Movie, NBA and University) and then applied to the held-out site. Like a Cross Validation experiments with each site. For CloseIE, they train only on in-vertical sites.  Level 3: Site-specific Knowledge, Level I & Level II are both zero-shot settings, Level III allows site-specific training data via weak supervision | | |
| **Evaluation/Testing Technique** | Comparision of proposed approach and baseline model were done.  The Precision, Recall and F1 Score were evaluated to calculate the accuracy of the defined model. | | |
| **Results** | <https://drive.google.com/file/d/1PqZ5nwXwplmPSLAoadHqGr3pERdJmouo/view?usp=sharing> | | |
| **Limitations/ Assumptions** | They sampled 100 error cases in each verticals from the unseen-vertical experiments and manually examined them to justy the error analysis | | |
| **Future Directions/ Open Issues** | - | | |