NiLinker summary / Review work ( unlinkable , not in candidate list )

Objective : What is NIL entity – In general domain clustering helps to resolve such issue

Fetching topK candidates from target kB that may represent

A new evaluation Ds – EvaNIL .

NIL entities are one which do not not find some candidate in KB conveying their meaning. not linkable.

So this thing affects recall / or when DS is not updated.

NIL entity is. Zero shot or Out of Bio Domain mention.?????

Zero shot can help identify out of domain mentions and eliminate them.

NIL linking is helpful for emergent entities which are not in DS ,

Besides, biomedical KBs miss Wikipedia-specific features. So challenging for KBio EL. Alos EvalNIL DS are not available.

Augment to REEL model , improvement visible.

Idea : already follow approach is fix threshold for matching disadvantage – semantic is lost since sent is lost in such case.

Emergent entity concept

--By generating KB candidates for the input NIL entity through a word-concept dictionary and then by applying a neural network using the attention model to determine the relevance of each KB candidate.

increase of 8%–9% in the number of new scientific articles from the end of the Second World War up until 2010. For instance, PubMed added 952,919 citations only in 2020.

which would allow the expansion of the target KB through the creation of a new subconcept of ‘‘Scientist” ie will link to closest with similar context & create subconcept.

Global & local concept of BioEL -

Model : possible approaches -> SVM ranker ( self feature extraction )

One possible approach is to match context and generate candidates, but the exact candidate entity is not present in KB , so context might be utilized , but that candidate must be related to a better candidate.

DilBert distance of false positives.

Major steps are string matching , hierarchical clustering , graph based .

Cluster entities sharing high similar entities .

Assign entity to random cluster

Graph appends semantic meaning & relation between NIL entities.

Typing is one approach which works – Noun phrase problem , if not linkable with noun phrase , classify it. to type one of 1339 . Find types which are more possible for linkable entities and if unlinkable types .

This approach for wiki data.

Other models try to propagate their semantic types to unlinkable entities through relations described in text, whilst our approach does the opposite, i.e, uses NIL entity linking to improve NEL.

Training : words pairing with sememes and assign weight to meaningful semems and generate embeddings.these combinations can generate meaning for all words in KB .So, weight was assigned to this semems generating embeddings for each semems.

So the objective is to select the most relevant KB concept.

Identification is a common method to give an empty candidate set for such an entity or below threshold score.

So an alternative approach for candidate set generation .

Hownet will not work, why ?

So KB derived Word concept dictionary .Key value -> word keys , concept values .

Tokenization of