

ABSTRACT:

Classifying biomedical literature is a difficult and challenging task, especially when a large number of biomedical articles should be organized into a hierarchical structure. In this paper, we present an approach for classifying a collection of biomedical text abstracts downloaded from Medline database with the help of ontology alignment. To accomplish our goal, we construct two types of hierarchies, the OHSUMED disease hierarchy and the Medline abstract disease hierarchies from the OHSUMED dataset and the Medline abstracts, respectively. Then, we enrich the OHSUMED disease hierarchy before adapting it to ontology alignment process for finding probable concepts or categories. Subsequently, we compute the cosine similarity between the vector in probable concepts (in the "enriched" OHSUMED disease hierarchy) and the vector in Medline abstract disease hierarchies. Finally, we assign category to the new Medline abstracts based on the similarity score. The results obtained from the experiments show the performance of our proposed approach for hierarchical classification is slightly better than the performance of the multi-class flat classification.