

Documentação do projeto

Grupo 7

Tomás Páscoa nº51592

João Vedor nº56311

Introdução

O tema deste projeto é “a simple ontology alignment framework in python: using lexical similarity and LLMs” e o projeto tem como objetivo desenvolver código em python que compare a semelhança entre as classes de duas ontologias previamente escolhidas.

Passos para concretizar o projeto

- 1b. Escolher 2 ontologias e importá-las no código
- 2b. Fazer parsing das classes das ontologias. Se uma classe tiver mais do que uma label escolhe-se a label que aparece com mais frequência dentro das labels associadas a uma determinada classe.
- 3b - Gerar BERT embeddings para as labels das classes depois de se fazer o parsing
- 4b - Percorrer as classes de ambas as ontologias e aplicar uma medida de string similarity (ex: levenshtein)
- 5b - Aplicar a medida de semelhança cosine para calcular a semelhança entre embeddings
- 6b - Calcular a média entre a semelhança léxica e a semelhança baseada nas embeddings para todos os pares de classes
- 7b - Mostrar os resultados para cada par

Testar com ontologias da conference e depois com as da anatomy

Gerar a matriz $m \times m \times 2$