NeighBert : Medical Entity Linking Using Relation Induced Dense Retrieval

Springer Nature ‘21

Problem Description:

Entity linking is the task of identifying mentions in documents and mapping them to their correct concept name in knowledge base.Knowledge Base is a huge collection(4.3 B) words in Biomedical Domain.

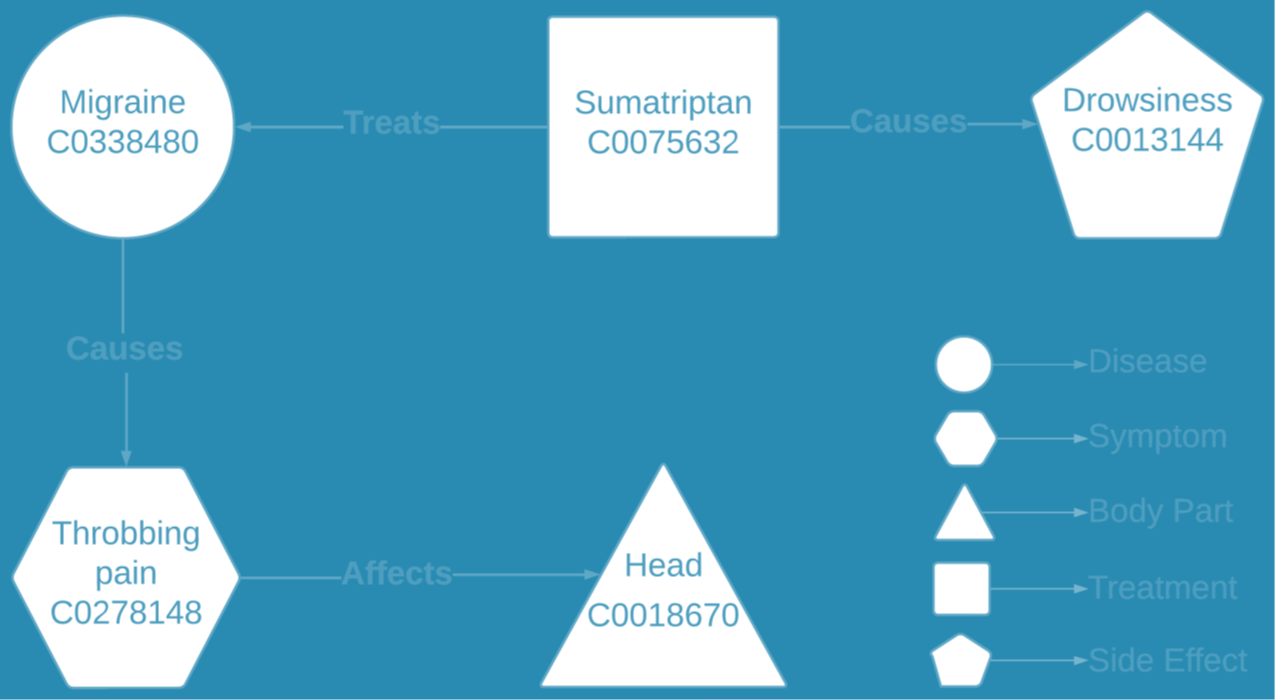
There are certain issues that follow up while finding an efficient mapping technique and that is resolved in entity linking tasks.

Proposed idea:

The idea is to Enrich the semantic encoding with structural relationships other than context encoding. KB here means entities and relationships that are already there in umls.NB is a pretrained model using Entities and relations present in UMLS.Context is used while Pre Training NB here and , mask for one of part of triplet is done.

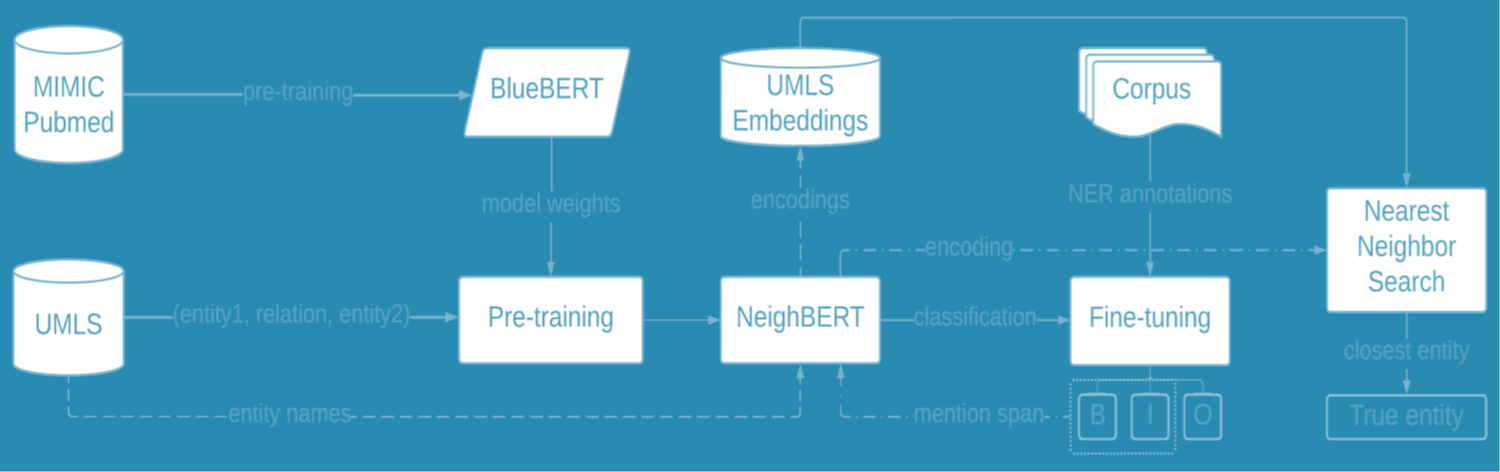
It is different from Normal GNN since it uses only model weights from BlueBert instead of some embedding which stores the structure of the graph .

An Example of structural relationships in medical concepts:



Model :

→NeighBERT does not require graph storage, instead updates model weights to learn how entities are related.

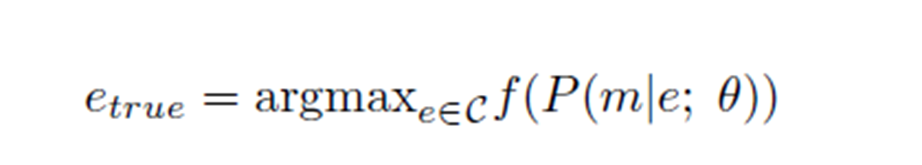


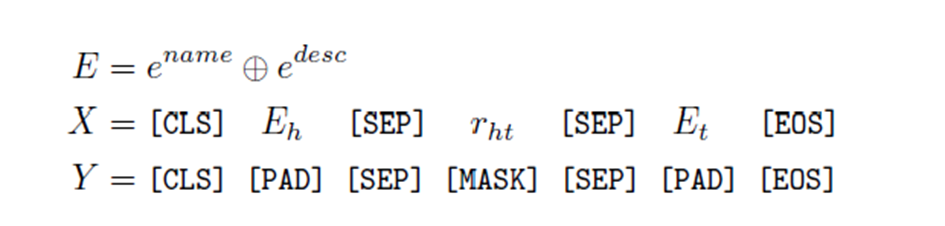
→Using model weight means to use weights to generate embeddings. Semantic + structural info both used.Drowsiness might not be covered using context.

→For pretraining MLM is used , Y as target class is used .Here ename and edesc are rep by E = xor of both .Objective is to learn a function with some parameter , maximizing probability for corr match.

→Not confirm whether model as a whole back propagates , since final decision is made by finding nearest neighbor in an index that is comprised of encoding of all the entities from the UMLS

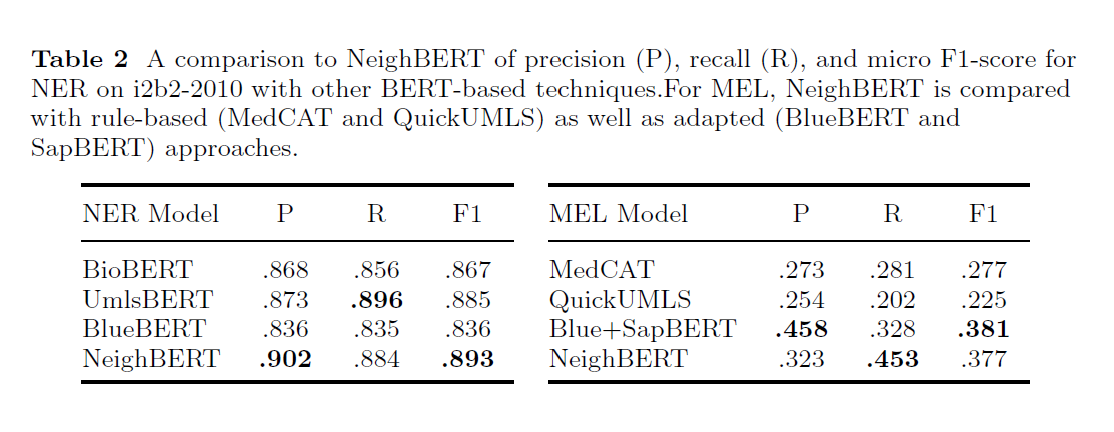
→A KB G à {entity , relation pairs } . Task is to learn function which maps relation ( mention à relation ) .





Training →

Comparison / Results →



Bottlenecks:

▪Difficulty detecting overlapping entities like two cancer names or sub mentions within mentions.

▪Only 60% mm concepts have definition in UMLS.