**LATTE: Latent Type Modeling for Biomedical Entity Linking**

Published → AAAI -2020

**Problem Description→**

* Effective Entity Linking task , same as previous paper.
* If the precise semantic type info is known for mentions , disambiguation becomes easier .
* Fine graining is the process of genrationg(unsupervised) semantic types other than that are already present in umls using known types during learning (known types used are umls types for now).
* IN UMLS , there are only 128 semantic types but 80k concepts so need for more specific type information can help in linking task.
* Example →the entities Type 2 Diabetes Mellitus and Parkinson Disease both have the semantic type Disease or Syndrome, but the former is a metabolic disorder, while the latter is a nervous system disorder. In this case, the finer-grained type can be the body system where the disease occur.

**Importance of problem→**

Accurate entity disambiguation is crucial to the understanding of biomedical context.

Many distinct biomedical concepts can have very similar mentions, and failure in disambiguation will lead to incorrect interpretation of the entire context. This will introduce

huge risks in medical-related decision making.

**Limitations of previous work→**

Mentions in bio domain are inherently ambiguous (not in the sense of actual ambiguency of language but a resolvable task to be acomplished).So , better the contextual representation ,better the resolution of ambiguency .

* If only surface level features are used it will bring ambiguity (enough context to be learned is not there to be learned ) .
* Also , on candidate side , semantic similarity can be high , it is also a problem(one to many mapping).
* Neural model used in web are hence not effective in this domain( due to long spans and inherent disambiguation in entities and mentions ).

**Intution of current work→**

Latte is neural network based model which genrates latent types for mentions and entities , based on their implicit attributes . which contribute in the simillarity score between mention and candidate to enhance the accurarcy of linking and disambiguation task. Latte does entity disambiguation along with type learning .

We jointly utillize the score generated on basis of interaction between mention and candidate and score from latent type generated similarity.

**Proposed Model** →

* Motivation is to model the latent fine-grained types for all the entities in the knowledge-base without direct supervision .
* Ground truths are the actual findings of fine graining which are not available with us , leading to treating fine graining as latent work of modelling.
* Constraints → Binary pairwise relation constraint & type hierarchy constraint( not clear why they are termed as constraints )

The type classifiers are sharing weights between them and The similarity of the two output latent type distributions is used as another mention-candidate similarity

score.

Layers of the model:

Superscripts p and c for mention and candidate sequences respectively,

mand n denote their corresponding sequence lengths

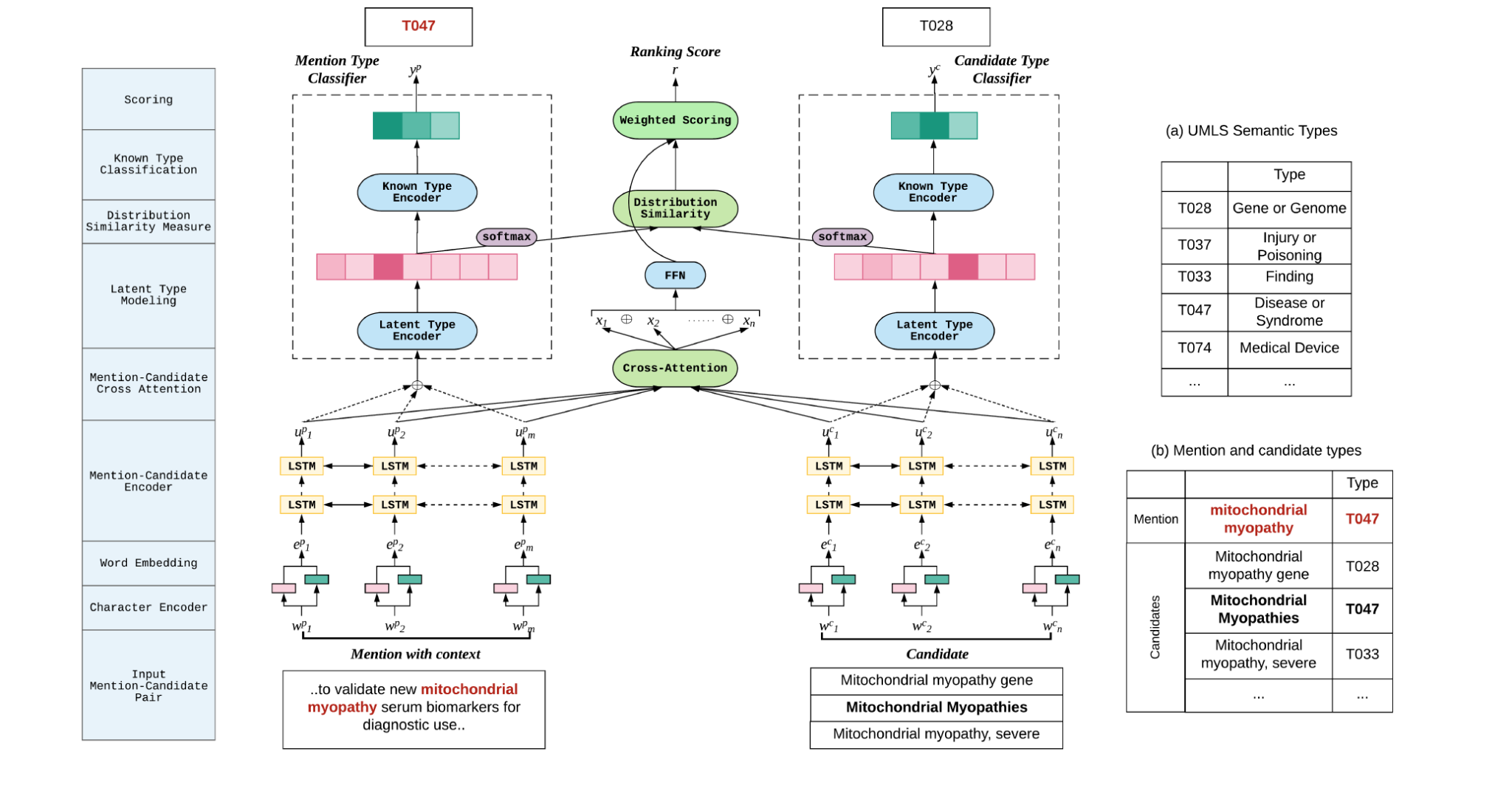
* Embedding Layer

This layer of model is responsible for generating embeddings for the input in form of mention sequence and candidate sequence .

I/P → word tokens as for the mention and the candidate sequences respectively

O/P→ , embeddings for each word token.

To resolve OOV words problem & utillize word level interaction , embedings are modified to first character level (using glove)interaction , concenated and passed through neural net then to get word embeddings as final o/p.(char embedding concenated , passed through nn then result is concenated with word embedding pre obt from glove)



* Encoder layer

→ I/P - output of embedding layer that is ei

O/p-contextual representation as  for the sequences.

* Cross attention layer →

I/P →

interaction bet Uip and Uic , Output is a matrix S which contains sij as score of interaction .

O/P →This sij is used to calculate attention mention to candidate and candidate to mention→in form of attended vector Xj , all attended vector concatenated to form X for

attention based relevence score between two sequences

(output of this layer is

Task 2→ takes op of encoder layer as ip , concatenates them for mention and candidates. Further passed over FFN and softmax to get probability distribution over k types ( to which it is more similar) .

Now this distribution are compared for G using cosine .

Type classifier are used which takes encoded representation to generate latent types.

**Experiments and Results** →

**Conclusion**

During training it requires types for minimizing loss to earn type generation

Dataset used → MedMention (made of UMLS concepts) ,de-identified corpus of dictated doctor’s notes (annotated with ICD concepts).

Input → mention phrase with words Wp1 wp2 etc .p , set of candidates c1 c2 c3 , model gives relevance score for each R(p,c). Word token for mention and candidate sequence respectively .

Input format might be similar to last paper and it was better explained there .

Embedding layer→ char to char to word (either

L1(char & word embedding ) → Lstm for contextual representation → cross attention layer record relevance between mention and candidate (o/p is S a matrix recording score of interaction between each pair m\*n entries ) → mention to candidate and candidate to mention score → attended vector to score relevance between sequences .

Sequence for latent entity type classifier → encoder se o/p → fixed dimen vector for mention and candidate → probability distribution over k latent types → **G score → cosine** similarity→how type is generated → supervised known type classifier (trained to predict entity type) takes represe and applies relu →

