Biomedical Interpretable Entity Representations

Garcia-Olano, D., Onoe, Y., Baldini, I., Ghosh, J., Wallace, B., Varshey, K. “Biomedical Interpretable Entity Representations”. Findings of the Association for Computational Linguistics (ACL-IJCNLP 2021)

Motivation

Entities over text = typically embedded in dense vector spaces with pre-trained language models (BERT,etc).

[0.519, 0.917, -0.935, 0.891, 0.396, 0.711, 0.479, 0.417, 0.744, -0.254,
-0.174, 0.233, -0.315, 0.497, -0.516, 0.22, -0.679, 0.389, -0.683, 0.909,
23, 0.528, 0.116, 0.334, 0.717, 0.857, -0.262, 0.624, -0.178, -0.045, -0.
-0.952, 0.4, 0.356, 0.091, 0.976, -0.337, -0.002, 0.054, 0.512, -0.312,
.278, -0.409, -0.655, -0.294, -0.453, 0.735, 0.461, 0.282, -0.43, -0.838,
3, -0.736, -0.001, 0.889, -0.228, 0.645, 0.883, 0.805]

[0.656, 0.407, 0.568, -0.035, -0.842, -0.257, 0.202, -0.31, 0.886, 0.386,
34, -0.823, -0.929, -0.068, -0.238, 0.236, -0.463, 0.56, -0.687, -0.521,
88, 0.54, 0.047, -0.434, -0.009, 0.59, 0.971, 0.798, 0.202, 0.225, 0.131,
88, 0.44, -0.835, -0.032, -0.935, 0.318, 0.72, -0.23, -0.903, 0.912, -0.8
0.981, -0.23, 0.797, -0.785, -0.583, 0.055, -0.511, 0.413, -0.757, 0.914,
943, 0.62, -0.78, 0.888, 0.288, 0.807, -0.207, -0.284]

Motivation

Entities over text = typically embedded in dense vector spaces with pre-trained language models (BERT,etc).

```
>>> word_embedding_for_happy
[0.519, 0.917, -0.935, 0.891, 0.396, 0.711, 0.479, 0.417, 0.744, -0.254,
-0.174, 0.233, -0.315, 0.497, -0.516, 0.22, -0.679, 0.389, -0.683, 0.909,
23, 0.528, 0.116, 0.334, 0.717, 0.857, -0.262, 0.624, -0.178, -0.045, -0.
 -0.952, 0.4, 0.356, 0.091, 0.976, -0.337, -0.002, 0.054, 0.512, -0.312,
.278, -0.409, -0.655, -0.294, -0.453, 0.735, 0.461, 0.282, -0.43, -0.838,
3, -0.736, -0.001, 0.889, -0.228, 0.645, 0.883, 0.805]
```

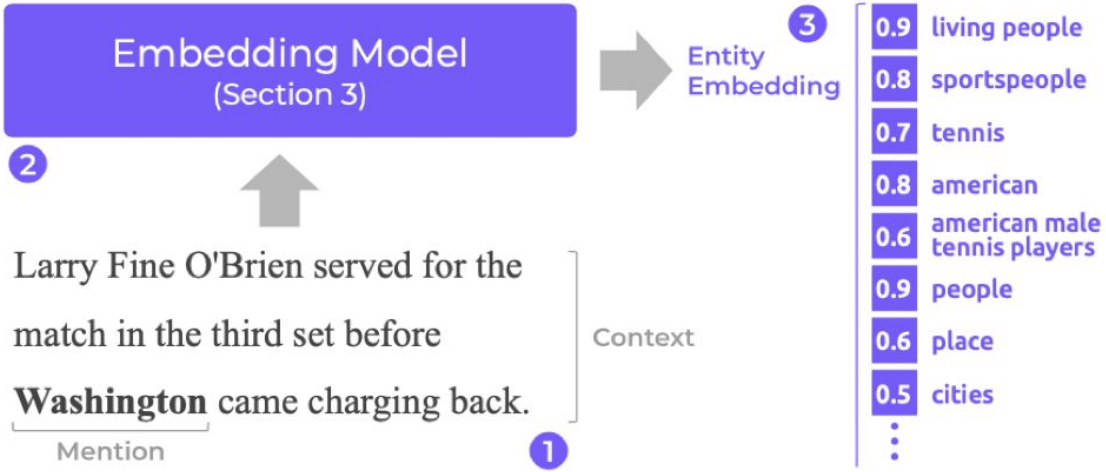
→

```
>>> word_embedding_for_sad
[0.656, 0.407, 0.568, -0.035, -0.842, -0.257, 0.202, -0.31, 0.886, 0.386,
34, -0.823, -0.929, -0.068, -0.238, 0.236, -0.463, 0.56, -0.687, -0.521,
88, 0.54, 0.047, -0.434, -0.009, 0.59, 0.971, 0.798, 0.202, 0.225, 0.131,
88, 0.44, -0.835, -0.032, -0.935, 0.318, 0.72, -0.23, -0.903, 0.912, -0.8
0.981, -0.23, 0.797, -0.785, -0.583, 0.055, -0.511, 0.413, -0.757, 0.914,
943, 0.62, -0.78, 0.888, 0.288, 0.807, -0.207, -0.284]
```

Not immediately interpretable.

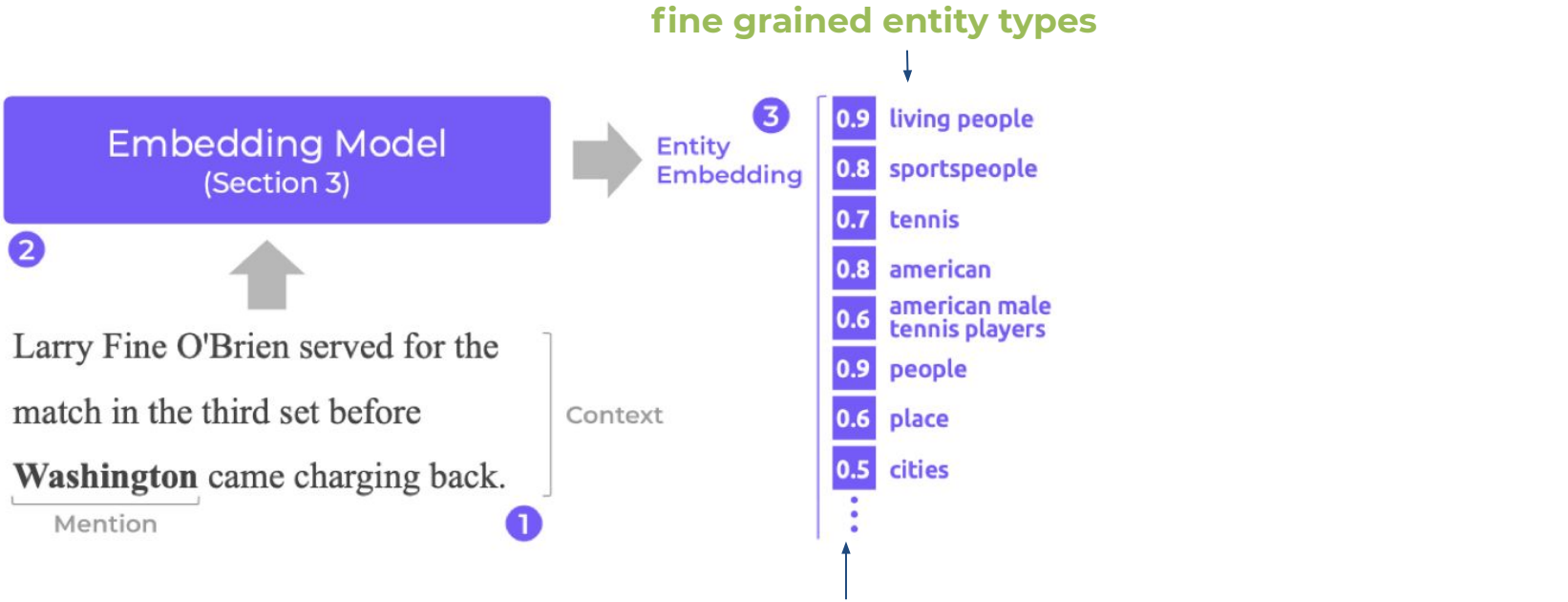
Dense Entity Embeddings = Give good performance for entity-related tasks, but using them in those tasks requires additional processing in neural models.

Onoe et al* learn human readable interpretable entity representations that achieve high performance without additional learning (“out of the box”)



“Interpretable Entity Representations Through Large Scale Typing”
Yasumasa Onoe & Greg Durrett . Findings of EMNLP 2020

Onoe et al* learn human readable interpretable entity representations that achieve high performance without additional learning ("out of the box")



experiments using Ultra Fine Entity Type system **(10k)**
and Wikipedia Categories Type System **(60k)**

Can we adapt IERs for the **Biomedical Domain?**

[Glesatinib] is a dual inhibitor of c-Met and SMO
that is under phase II clinical trial for non-small cell lung cancer.

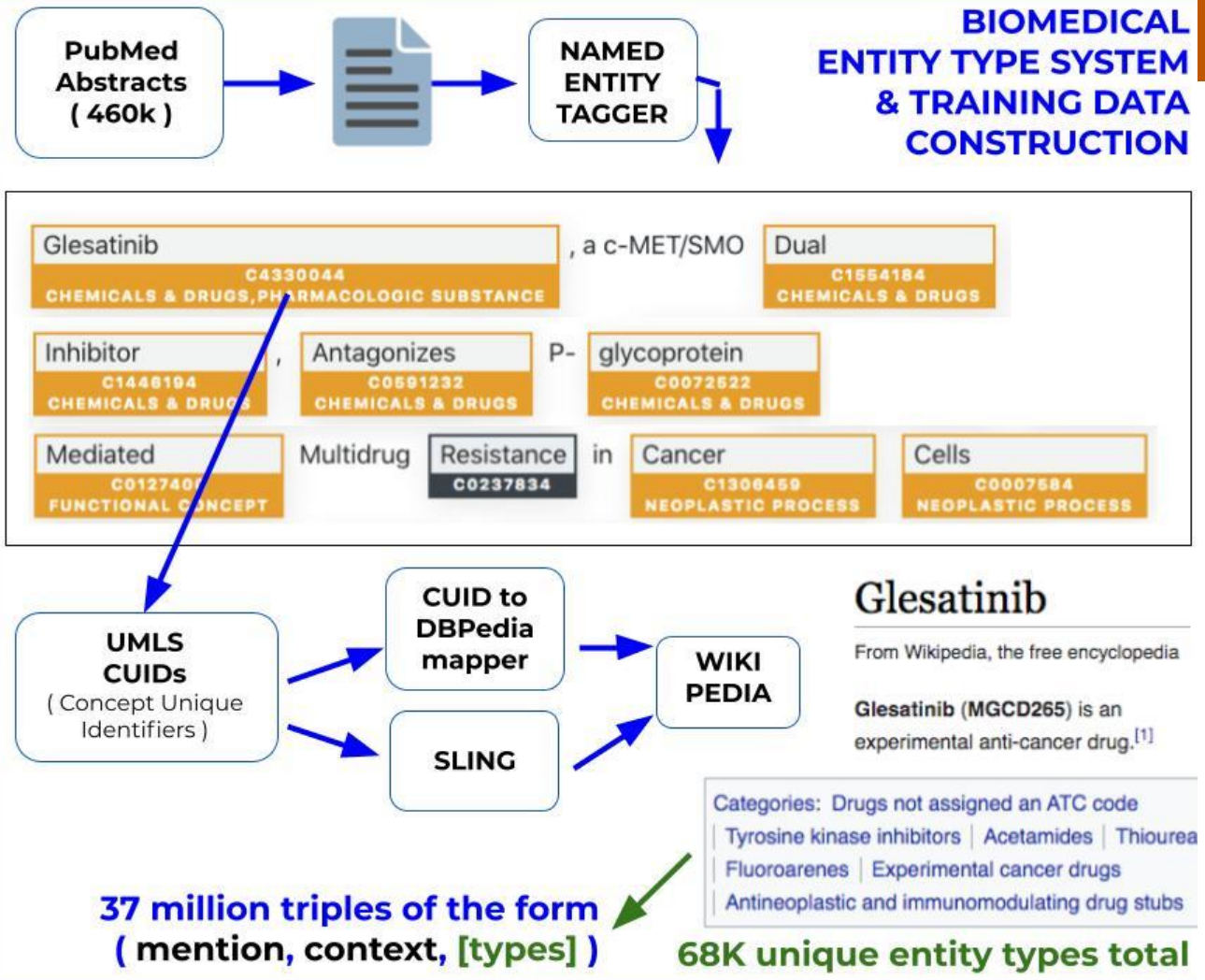
Can we adapt IERs for the **Biomedical Domain?**

[Glesatinib] is a dual inhibitor of c-Met and SMO
that is under phase II clinical trial for non-small cell lung cancer.

Most probable entity types for mention/context

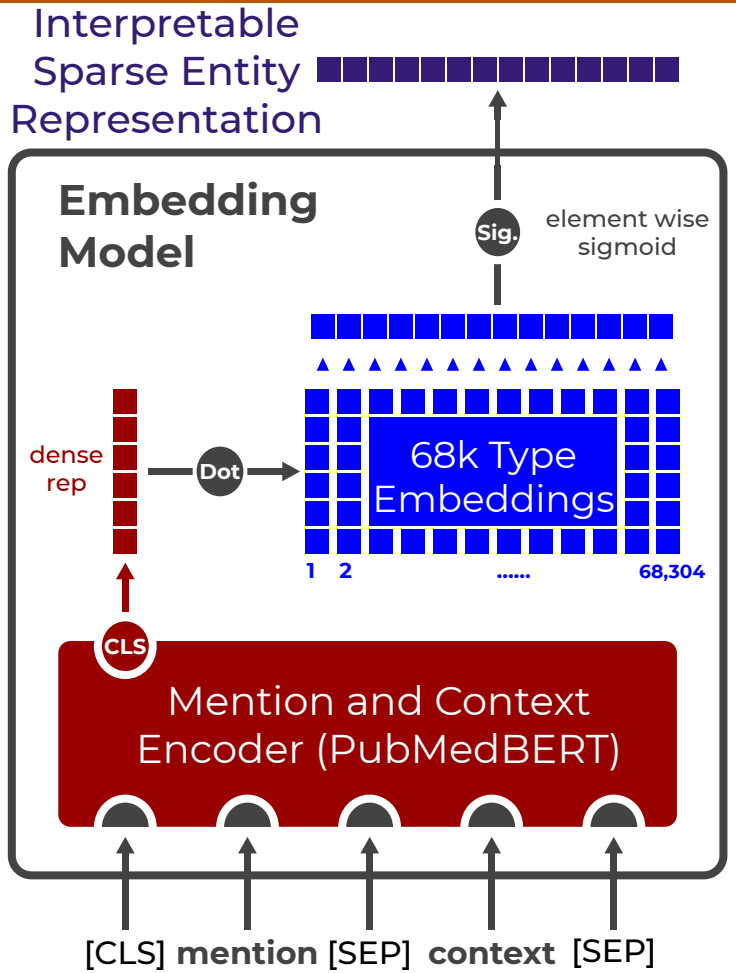
world health organization essential medicines	0.4941
pyridines	0.4073
diols	0.3539
cancer treatments	0.3260
carboxylate esters	0.2376
chloroarenes	0.1984
rtt	0.1879
hormonal antineoplastic drugs	0.1768
antineoplastic drugs	0.1037
alcohols	0.0771
prodrugs	0.0315
peptides	0.0300
methyl esters	0.0223
merck	0.0191
transgender and medicine	0.0135
teratogens	0.0130
world anti-doping agency prohibited substances	0.0124
peripherally selective drugs	0.0103
human proteins	0.0099
ureas	0.0090
withdrawn drugs	0.0089
iarc group 2a carcinogens	0.0073
prostate cancer	0.0066
mechanisms	0.0066
chemotherapy	0.0058
aromatase inhibitors	0.0057

of 60k wiki entity types

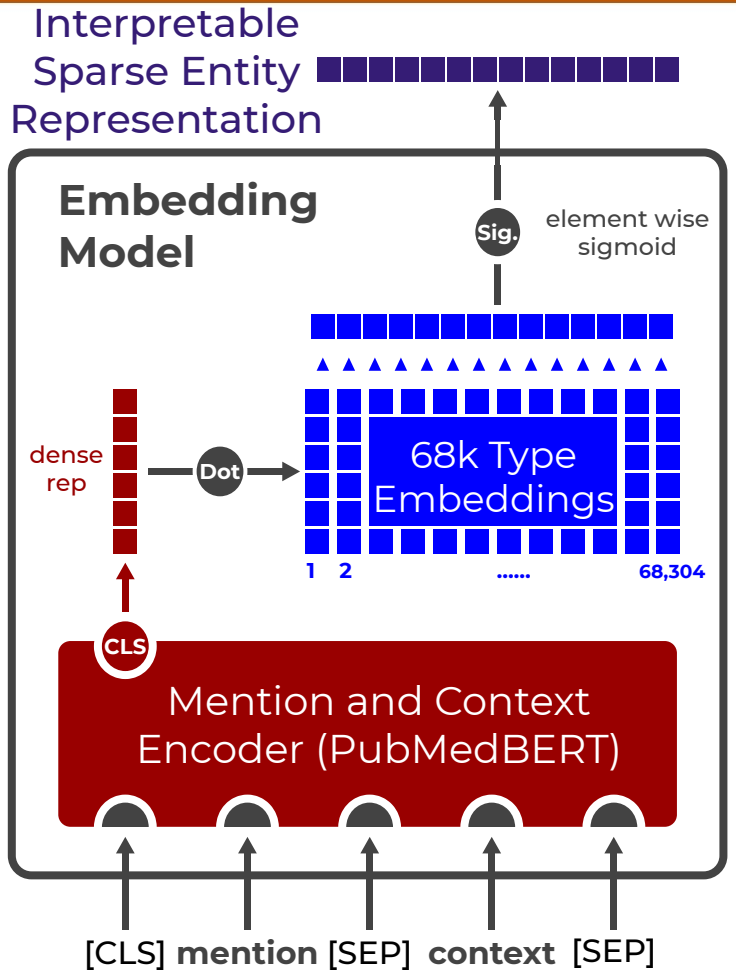


Distant Supervision to **construct** **Entity Type System** and **Training Data**.

Training Biomedical IERs



Training Biomedical IERs



Training loss:

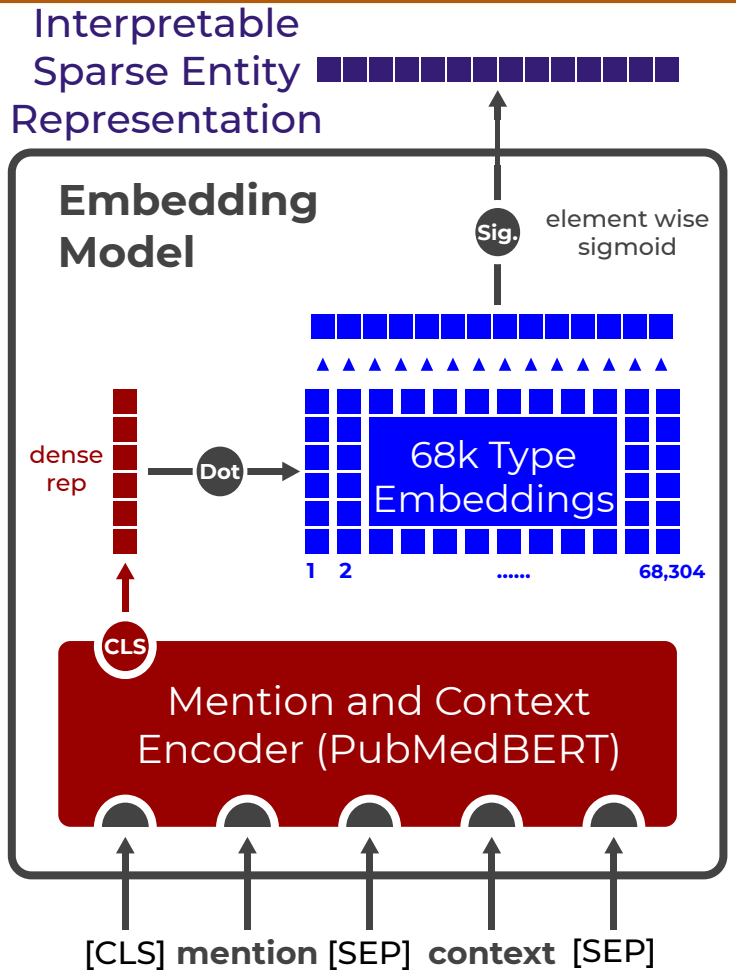
Independent sum
of binary cross entropy losses
over all all entity types T
over all training examples D.

$$-\sum_{i}^{D}\sum_{j}^{T} t_{ij}^{*} \cdot \log(t_{ij}) + (1 - t_{ij}^{*}) \cdot \log(1 - t_{ij}),$$

true types for training instance i

predicted types

Inference using Biomedical IERs



Inference via simple cosine similarity between Biomedical IERs without fine-tuning on task data !

(1) **Named Entity Disambiguation** (NED) on Clinical Entities.

Given a entity mention, context & set of candidate entities
identify which of the candidates is the true one linked to the mention.

Model	Test Acc.	
	Dot Prod	Cosine Sim
BIER-PubMedBERT (ours)	80.1	**84.0**
BIER-SciBERT (ours)	76.4	77.3
BIER-BioBERT (ours)	71.9	75.9
Onoe and Durrett (2020)	63.6	69.8
Popular Prior	73.9	-
PubMedBERT (Gu et al., 2020)	77.6	-
SciBERT (Beltagy et al., 2019)	77.4	-
BioBERT (Lee et al., 2019)	77.9	-

Prior work

Fine tuned approaches

Table 2: BIER zero shot test results vs Logistic Regression Baselines trained on task data for NED task

(2) Entity label Classification for Cancer Genetics

Model	Test Acc.			
	L2 Dist		Dot Prod	
	Dense	Sparse	Dense	Sparse
BIER-PubMedBERT	85.5	86.8	**88.2**	**87.5**
BIER-SciBERT	70.8	77.0	72.8	76.8
BIER-BioBERT	83.4	85.9	85.6	86.8
Onoe and Durrett (2020)	63.9	55.1	60.0	59.9
PubMedBERT	77.3	-	69.3	-
SciBERT	74.4	-	75.2	-
BioBERT	67.6	-	59.6	-

Prior work

Fine tuned approaches

Table 3: Test accuracy on Cancer Genetics data using a nearest neighbor classifier (k=1) without fine-tuning based on sparse output or intermediate dense embeddings using L2 or Dot Product distance metrics.

(2) Entity label Classification for Cancer Genetics

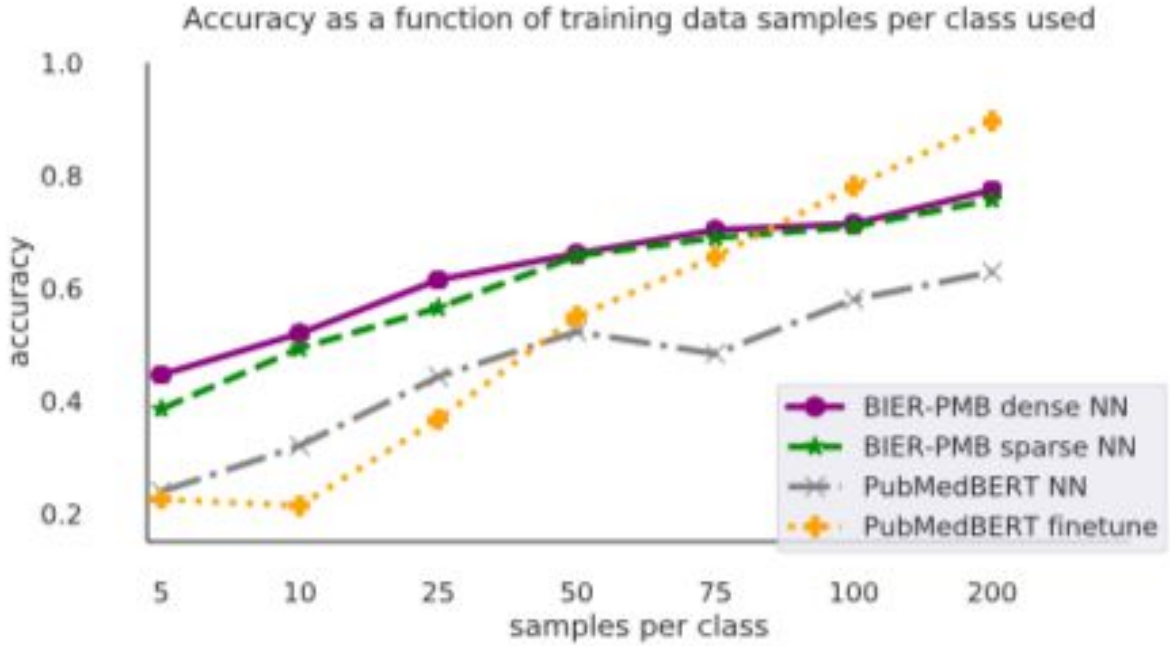
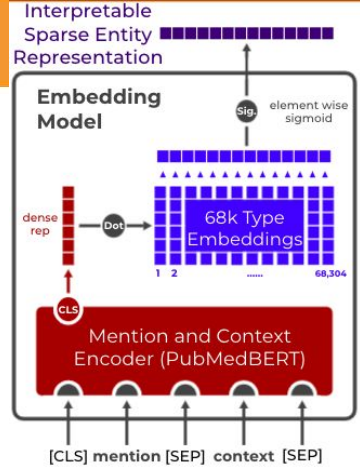


Figure 3: Results for the entity label classification task under varying amounts of supervision.

Debugging with BIERs

Allows for error analysis at the component level to identify areas lacking in supervision and/or possible changes to the type system.

Debugging with BIERs



Allows for error analysis at the component level
to identify areas lacking in supervision
and/or possible changes to the type system.

How well the model could have done
had it known to fallback to
using the intermediate dense embedding
in cases where the sparse representation
led to an **incorrect prediction**

	Test Acc.			
Task	Dense	Sparse	Combined	Δ
NED	84.0	81.0	**91.7**	+7.7
ELC	87.5	88.2	**91.9**	+3.7

Table 5: Results for both tasks showing improvements that could have been achieved by combining intermediate dense and interpretable sparse output embeddings generated by the same BIER-PubMedBERT model.

Motivation for **future work**
on developing a dynamic approach
to making predictions
that is a function of model confidence.

Thank you!

Code and data available:

https://github.com/diegoolano/biomedical_interpretable_entity_representations

IBM Research
Science for Social Good

TEXAS
The University of Texas at Austin

LVX VERITAS VIRTVS
Northeastern University