One-shot to Weakly-Supervised Relation Classification using Language Models

Thy Thy Tran, **Phong Le**, Sophia Ananiadou

Contact: thytran@ukp.informatik.tu-darmstadt.de
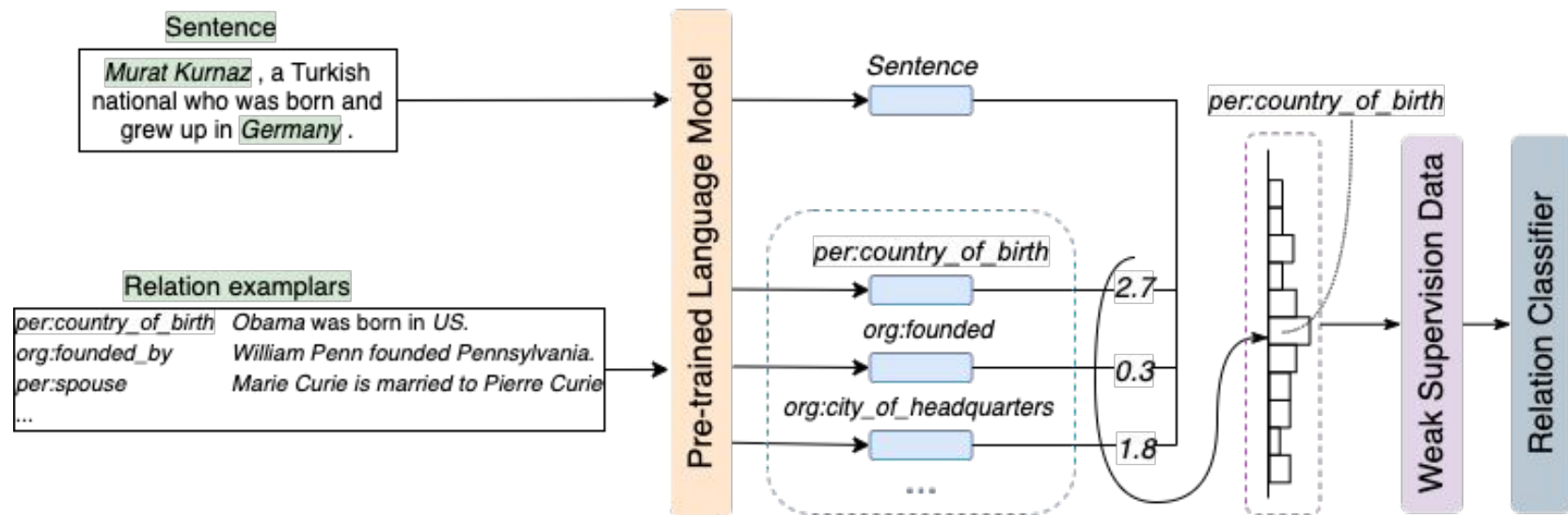
Code: https://github.com/ttthy/noelaw

Takeaways

- LMs can be used to generate noisy data for relation classification

- Noise reduction mechanisms can improve relation classification on substantially

  noisy data

Weak Supervision from Language Models

1. Assign to each relation type a very simple exemplar
2. Given a raw sentence, compute the match scores between it with the exemplars. The *best match* is the *noisy label* of the raw sentence.
3. Train a relation classifier on the resulting data

<u>No</u>isy Chann<u>el</u> <u>A</u>uto-encoder (NoelA)

Using entity type pairs as inductive biases to eliminate unrealistic cases e.g. "US was born in Europe"

Employing a noisy channel to tolerate the annotation noise
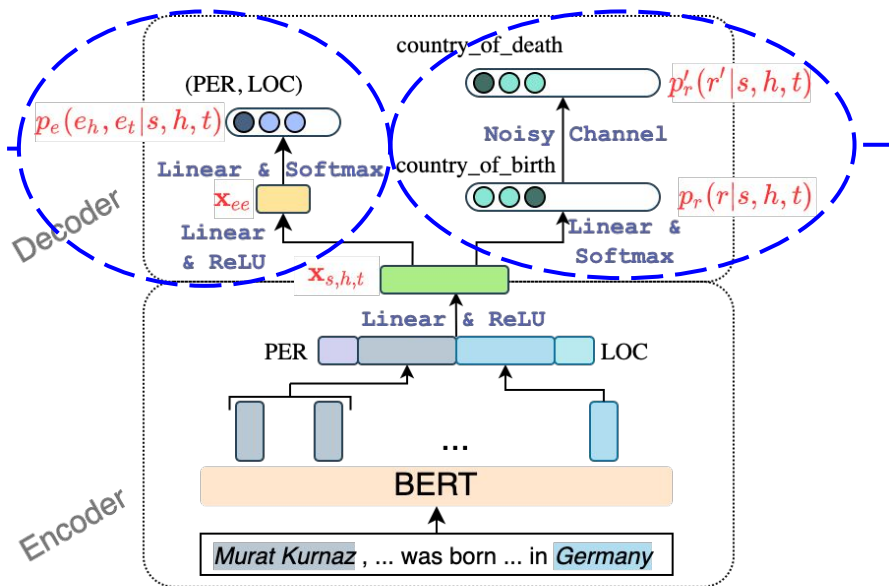
Results

	TACRED		reWiki80	
	Acc. (%)	Abs.+	Acc. (%)	Abs.+
Matching				
Random	2.44	-	1.25	-
Frequency	15.04	-	1.25	-
Pretrained Language Models				
GPT2-small	0.27	-	1.73	-
SpanBERT-base	8.36	-	6.45	-
BERT-base	15.46	-	27.48	-
Noisy Data				
Bootstrap-hard	19.28 ±0.42	3.82	29.76 ±0.16	2.28
NoelA	**24.79** ±0.68	9.33	**33.17** ±0.39	5.69
–ETR	21.54 ±0.69	6.08	32.48 ±0.67	5.00
–DR	21.28 ±0.54	5.82	32.65 ±0.11	5.17
–NC (BERTwET)	19.03 ±0.34	3.57	30.06 ±0.14	2.58
Gold Data				
BERTwET (sup.)	82.73 ±0.99	67.27	73.92 ±3.46	46.44

Substantial improvements from using BERT for one-short (similar to "matching the blanks" by Soares et al. 2019)