

Project

Requirements

- ▶ You are going to explore something new (could be reproduction study)
- ▶ Max 5 pages ACL style files:
<https://github.com/acl-org/acl-style-files>
- ▶ Bibliography and appendix do not count
- ▶ Result should be close to a publishable paper!

Requirements

- ▶ You should focus on Named Entity Recognition (NER) or Relation Extraction (RE)
- ▶ You can pick one of the topics proposed by us, or propose your own
- ▶ If you plan to propose your own topic, get in touch with Rob
- ▶ Every group is required to implement a Baseline for NER

Requirements


- ▶ More advanced than most-frequent class (O)
- ▶ Could be viterbi, CNN, LSTM, BERT, etc.
- ▶ For one of the projects, a machine learning classifier could be the best fit.

Requirements

NER task:

	Barack		Obama		was		born		in		Hawaii	
	B-PER		I-PER		O		O		O		B-LOC	

Note that entities can be nested:

				
Wozniacki	vinder	iik	australian	open
Wozniacki	vinder	ikke	Australian	Open

Requirements

IT	University	of	Copenhagen	
B-ORG	I-ORG	I-ORG	I-ORG	
0	0	0	B-LOC	

Datasets

- ▶ EWT: <https://aclanthology.org/2021.findings-acl.158.pdf>
- ▶ Tweebank: <https://aclanthology.org/2022.lrec-1.780.pdf>
- ▶ DaN+: <https://aclanthology.org/2020.coling-main.583.pdf>
- ▶ CrossNER: <https://ojs.aaai.org/index.php/AAAI/article/view/17587/17394>
- ▶ WikiAnn, MultiCoNER II: Use with caution (silver data!)

Datasets

Varieties:

Name	Langs	Domains	Nested	Human	# labels
EWT	EN	5 web	Y	Y	12
Tweebank	EN	Social	N	Y	4
DaN+	DA	News,Social	Y	Y	12
CrossNER	EN	5 wiki	N	Y	39
WikiAnn	295	wiki	N	N	3
MultiCoNER II	12	wiki	N	N	33

Datasets

If these are insufficient for your research question: find another or make one!

Project ideas

Project ideas

Can we improve NER systems?

- ▶ Might be hard
- ▶ Find challenging situations to focus on
- ▶ Incorporate POS tags?

Project ideas

Checklist evaluation

- ▶ Checklists have been proposed to do controlled evaluation of classification systems:

B Testing NER with INV Same pred. (inv) after removals / additions			
@AmericanAir thank you we got on a different flight to [Chicago → Dallas].	inv	pos neutral	x
@VirginAmerica I can't lose my luggage, moving to [Brazil → Turkey] soon, ugh.	inv	neutral neg	x
...			
Failure rate = 20.8%			

- ▶ <https://aclanthology.org/2020.acl-main.442.pdf>

Project ideas

Checklist evaluation

- ▶ Checklists have been proposed to do controlled evaluation of classification systems:

B Testing NER with INV Same pred. (inv) after removals / additions			
@AmericanAir thank you we got on a different flight to [Chicago → Dallas].	inv	pos neutral	x
@VirginAmerica I can't lose my luggage, moving to [Brazil → Turkey] soon, ugh.	inv	neutral neg	x
...			
Failure rate = 20.8%			

- ▶ <https://aclanthology.org/2020.acl-main.442.pdf>
- ▶ Can they be used for span-based tasks?

Project ideas

Cross-dataset NER (language/domain)

- ▶ How to exploit the strengths of multiple datasets for NER?
- ▶ They have different amounts of quality and different annotation guidelines

Project ideas

NER for new language/domain

- ▶ I.E.: Can we build an NE tagger for Frisian social media data?
- ▶ Annotate a small dev/test set, transfer from other train set(s)
- ▶ Evaluate existing model(s)

Project ideas

Quality check of silver (or gold) data

- ▶ Re-annotate existing data
- ▶ Find difficult cases for humans, do they overlap with the difficult cases for automatic systems?

Project ideas

Emerging NEs

- ▶ Most datasets have overlapping entities between train and test splits
- ▶ Performance is much lower if these are separated (40%)
- ▶ How can we overcome this gap?
 - ▶ gazeteers
 - ▶ data augmentation
 - ▶ pre-training?
 - ▶ self-training?

Project ideas

Unsupervised NER

- ▶ Remove the assumption of training data
- ▶ Start with easy cases, bootstrap from there
- ▶ Capitalization or embeddings spaces can be used for example

Project ideas

Active learning for spans

- ▶ Active learning = human-in-the-loop
- ▶ Automatically select new batch of training samples to annotate after every N annotations.
- ▶ <https://aclanthology.org/W15-1511.pdf>
- ▶ Propose to use SVM for POS tagging, and focus only on ambiguous cases.
- ▶ For NER, the ambiguity problem is less important, but the labels have to be in the correct order!
 - ▶ I would suggest to ignore BIO labels, and add them post-hoc

Project ideas

How to do nested NER

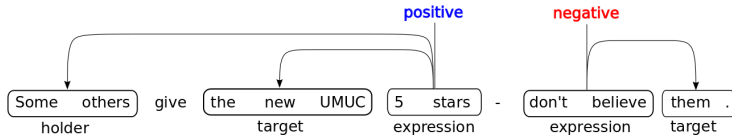
- ▶ Nested NER has a different structure as a task
- ▶ The second layer has much lower scores (<https://aclanthology.org/2020.coling-main.583.pdf> table 5)
- ▶ Hard to improve, as they are scarce!

Project ideas

For the ambitious: Relation Extraction

Project ideas

Relation extraction as sequence labeling



Project ideas

Pipeline versus sequence labeling

- ▶ Identify spans (span-labeling)
- ▶ Binary identification of relations
- ▶ Labeling of relations

vs.

- ▶ Convert task to sequence labeling

Project ideas

