Named Entity Recognition (NER)

Sequence models, Features for NER, Precision, Recall, F1, Deep Learning Models for NER

Mingyang Xu, mxu5@ncsu.edu
PhD Student in Dr. Samatova's research lab,
Department of Computer Science
North Carolina State University

Why Named Entity Recognition (NER)?

Motivation

- Unstructured text data is ubiquitous: news, blogs, etc.
- It's not feasible for human beings to process this data to identify important information
- Named entities can help us summarize the text data
- For question answering, answers are often named entities

Tasks of Named Entity Recognition (NER)

- Extract names
- Classify the extracted names

WWDC included a first-day keynote where Apple introduced upcoming software and new hardware. The 2017 conference was held on June 5 in San Jose, California. Tickets were priced at $1,599 and were distributed through a lottery system.

Potential Tags
LOCATION
ORGANIZATION
DATE
MONEY
PERSON
TIME

Traditional Approaches for NER

Supervised Methods

- Learn a model from annotated training examples for labeling new examples
  - Hidden Markov Model (HMM)
  - Maximum Entropy Models (ME)
  - Support Vector Machines (SVM)
  - Conditional Random Fields (CRF)

- Example of features used:
  - Current/previous/next word (word context)
  - Part-of-speech/Chunk tags
  - Previous/next word's label (label context)
  - Prefixes/suffixes

Evaluation for NER

The performance of NER [5] is usually measured using F1 score.

- $F_1 = 2 * \frac{precision * recall}{precision + recall}$
- $precision = \frac{\#\text{correctly predicted named entities}}{\#\text{ detected named entities}}$
- $recall = \frac{\#\text{correctly predicted named entities}}{\#\text{ present named entities}}$

Please note that the precision and recall are computed based on entities instead of tokens

Recent Works of NER

TwiNER: named entity recognition in targeted twitter stream – TwiNER [2]
A novel unsupervised NER system for targeted Twitter stream. In the first step, it leverages on the global context obtained from Wikipedia and Web N-Gram corpus to partition tweets into valid segments (phrases). In the second step, TwiNER constructs a random walk model for detecting named entities.

Joint Named Entity Recognition and Disambiguation – JERL [3]
A joint entity recognition and linking model, which jointly models NER and linking tasks and captures the mutual dependency between them. This model allows the information from the task of NER and entity linking to improve the performance of the other.

Boosting Named Entity Recognition with Neural Character Embeddings – CharWNN [4]
A language-independent NER system that uses automatically learned features only. Specifically, the system is based on the CharWNN deep neural network, which uses word-level and character-level representations (embeddings) to perform sequential classification.

Recent Works of NER

Named Entity Recognition with Bidirectional LSTM-CNNs – LSTM-CNNs [7]
A novel neural network architecture that automatically detects word- and character-level embeddings using a hybrid bidirectional LSTM and CNN architecture, eliminating the need for most feature engineering.

Neural Architectures for Named Entity Recognition – LSTM-CRFs [8]
A partially supervised topic model, which discovers the latent topics within each label, as well as unlabeled, corpus-wide latent topics. Specifically, for each document, PLDA introduces a set of latent topics within each label of the document and a set of latent topics without any labels.

Fast and Accurate Entity Recognition with Iterated Dilated Convolutions – ID-CNNs [1]
This paper proposes a faster alternative to Bi-LSTMs for NER: Iterated Dilated Convolutional Neural Networks (ID-CNNs), which have better capacity than traditional CNNs for large context and structured prediction.

Topic Modeling Software

Methods	Link	Language
Traditional Methods	OpenNLP	Java/Scala/R
	NLTK	Python
	Stanford NER	Java
	spaCy	Python
LSTM-CRFs	lstm-crf	Python
Stack-LSTM NER	stack-lstm-ner	Python
LSTM-CNNs	lstm-cnn	Python
ID-CNNs	id-cnn	Python
NER-LSTM	ner-lstm	Python
BiLSTM-CRF	bilstm-crf	Python

References

[1] Strubell, Emma, et al. "Fast and accurate entity recognition with iterated dilated convolutions." *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 2017.

[2] Li, Chenliang, et al. "Twiner: named entity recognition in targeted twitter stream." *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2012.

[3] Luo, Gang, et al. "Joint entity recognition and disambiguation." *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. 2015.

[4] Santos, Cicero Nogueira dos, and Victor Guimaraes. "Boosting named entity recognition with neural character embeddings." *arXiv preprint arXiv:1505.05008* (2015).

[5] Tjong Kim Sang, Erik F., and Fien De Meulder. "Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition." *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4*. Association for Computational Linguistics, 2003.

[6] Church, K. W. and Hanks, P. (1989). Word association norms, mutual information, and lexicography. In ACL-89, Vancouver, B.C., pp. 76–83.

[7] Chiu, Jason PC, and Eric Nichols. "Named entity recognition with bidirectional LSTM-CNNs." *arXiv preprint arXiv:1511.08308* (2015).

[8] Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., & Dyer, C. (2016). Neural architectures for named entity recognition. *arXiv preprint arXiv:1603.01360*.

8