Comparing Conditional Random Fields and LSTM Networks for Named Entity Recognition

Josef Gugglberger (student) Clemens Sauerwein (supvervisor)

December 18, 2019

LV 703605-6 Masterseminar

universität innsbruck

Motivation

test

Overview

1. Background & Related Work
2. Implementation Details
3. Evaluation and Comparison
4. Conclusion

Background & Related Work

Named Entity Recognition

Definition: NER

Named Entity Recognition is the task of locating and classifying named entities in unstructured text. A named entity is classified into a predefined set of categories.

Named Entity Recognition

Definition: NER

Named Entity Recognition is the task of locating and classifying named entities in unstructured text. A named entity is classified into a predefined set of categories.

James visited the Eiffel Tower in 2012.

↓

James [PERSON] visited the Eiffel [LOCATION] Tower [LOCATION] in 2012 [TIME].

Conditional Random Fields

Definition: CRF

A *Conditional Random Field* is a discriminative probabilistic classifier. It makes its prediction not just based on the input sample, but also based on the context of the input sample.

Conditional Random Fields

Definition: CRF

A *Conditional Random Field* is a discriminative probabilistic classifier. It makes its prediction not just based on the input sample, but also based on the context of the input sample.

$$p(y|x) = \frac{1}{Z(x)} \prod_{t=1}^{T} exp(\sum_{k=1}^{K} \theta_k f_k(y_t, y_{t-1}, x_t)) \tag{1}$$

where $Z(x)$ is an normalization function:

$$Z(x) = \sum_{y} \prod_{t=1}^{T} exp(\sum_{k=1}^{K} \theta_k f_k(y_t, y_{t-1}, x_t)) \tag{2}$$

Recurrent Neural Networks

Definition: RNN

RNNs are a special type of artificial neural networks that have a feedback loop feeding the hidden layers back into themselves. The loop provides a kind of memory that allow the network to better recognize patterns.

Recurrent Neural Networks

Definition: RNN

RNNs are a special type of artificial neural networks that have a feedback loop feeding the hidden layers back into themselves. The loop provides a kind of memory that allow the network to better recognize patterns.

- Suited for sequence labeling
- Problems with long term dependencies
- Vanishing and exploding gradient

Long-Short-Term-Memory Networks

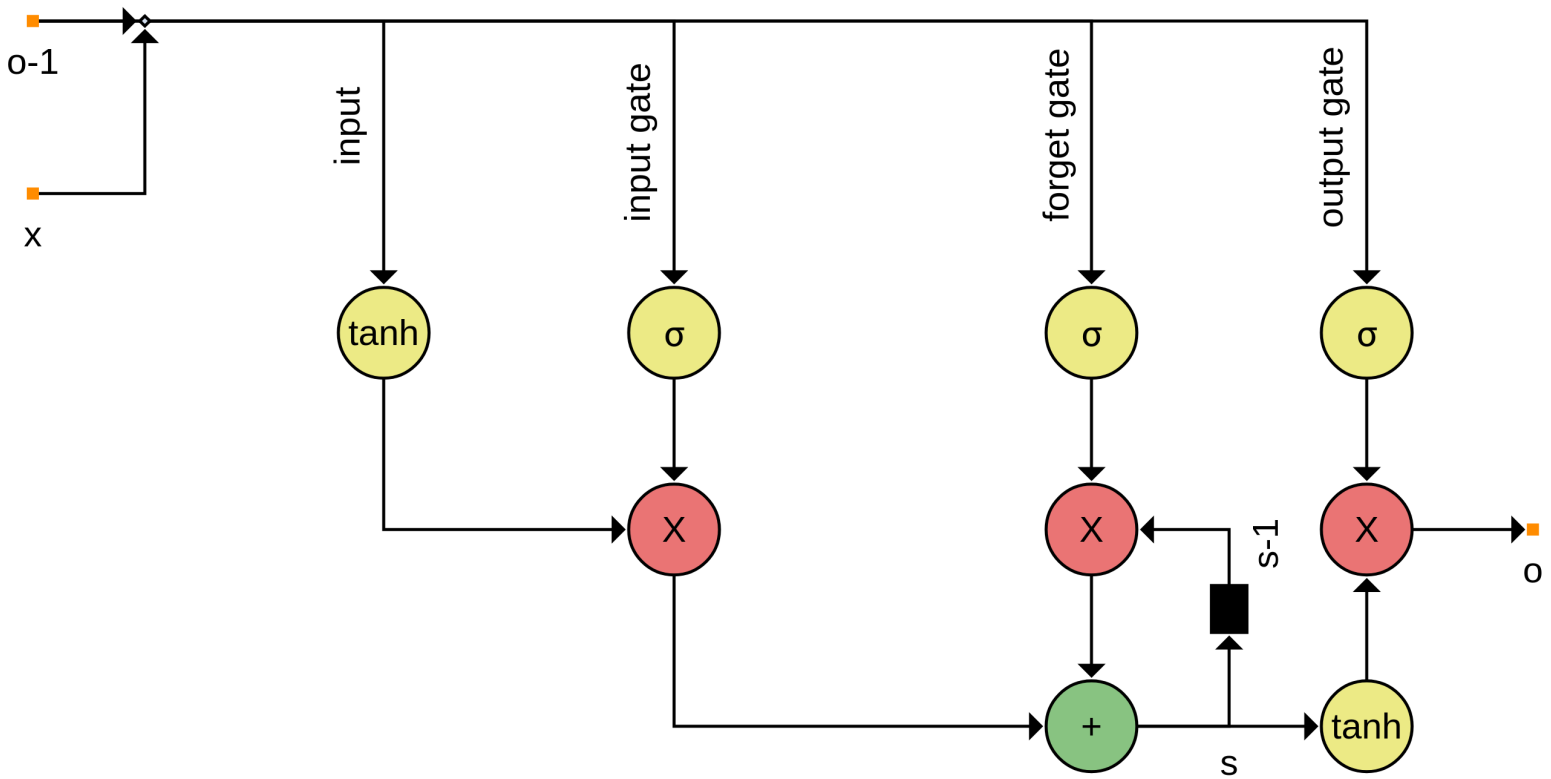
Definition: LSTM networks

LSTM networks are a special case of RNNs, which where designed to overcome issues with the vanishing gradient on long term relationships.

Long-Short-Term-Memory Networks

Definition: LSTM networks

LSTM networks are a special case of RNNs, which where designed to overcome issues with the vanishing gradient on long term relationships.



Implementation Details

Dataset

Conference on Computational NL Learning

CoNLL 2003 was a shared task on language independent named entity recognition.

Four types of Named Entities:

- Person
- Location
- Organization
- Miscellaneous

Dataset

Word	POS	Syntax Chunk	NE
U.N.	NNP	I-NP	I-ORG
official	NN	I-NP	O
Ekeus	NNP	I-NP	I-PER
heads	VBZ	I-VP	O
for	IN	I-PP	O
Baghdad	NNP	I-NP	I-LOC
.	.	O	O

Conditional Random Fields

Library:

- pycrfsuite

Features should describe characteristics of named entities.

- Word Features
- Sentence & Collection Features
- Dictionary Features
- Features from unsupervised ML algorithms

Word Features

- length of word
- the word starts with an upper-case letter
- the word contains an upper-case letter
- the word contains a digit
- the word contains a special character (-, /, etc.)
- word shape: 'Word' $\rightarrow$ 'Aa+', 'WORD' $\rightarrow$ 'A+', '2019-12-12' $\rightarrow$ '9999#99#99'

Sentence & Collection Features

- position of word in sentence
- number of occurrences in collection

Dictionary Features

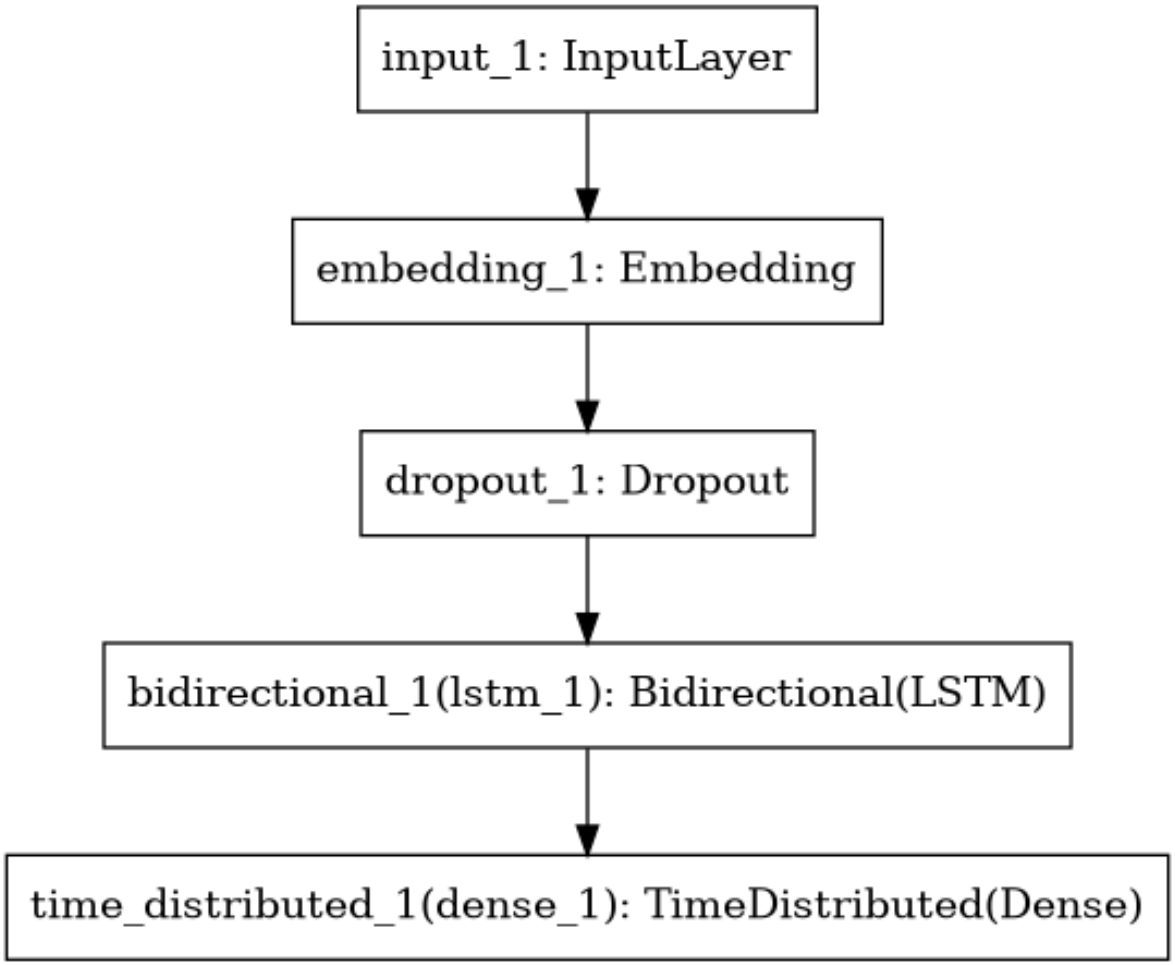
The word is contained in:

- stop-words list
- name list
- word list
- wordnet

Features from unsupervised ML algorithms

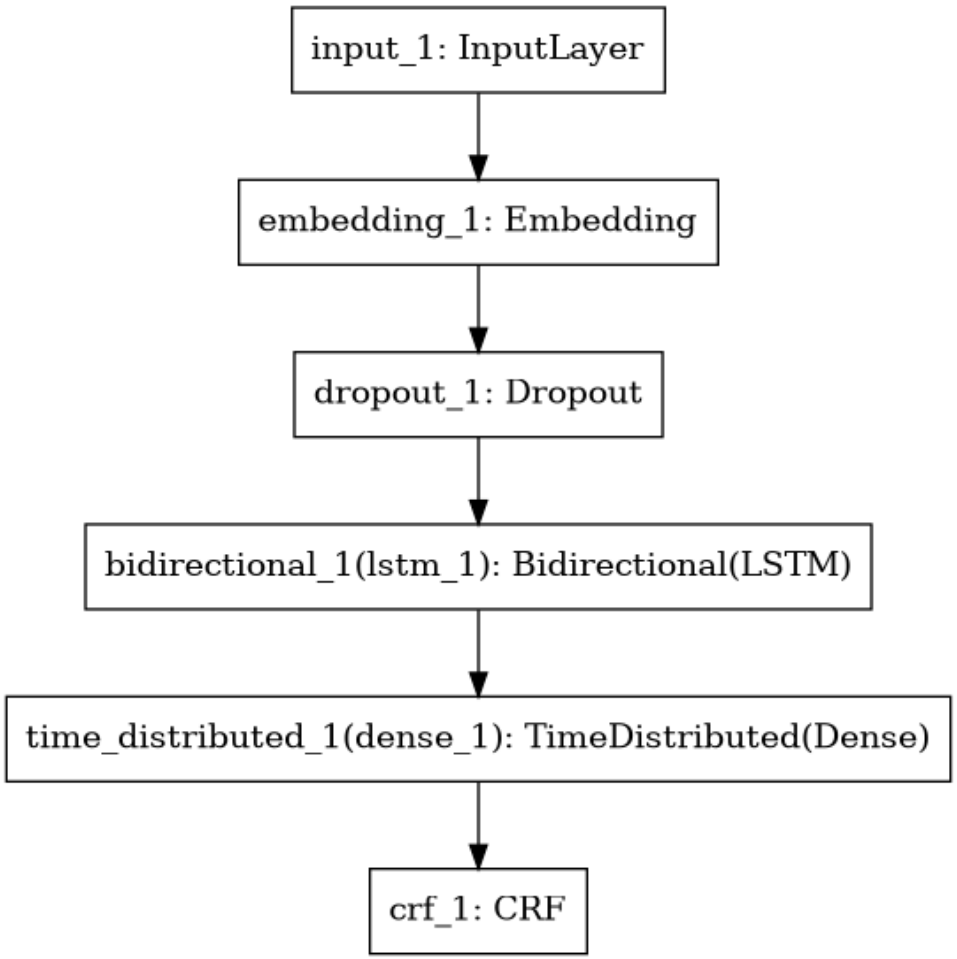
- brown cluster
- LDA topic
- w2c cluster

LSTM Network



The best of both worlds?

Combine the LSTM approach with CRF by adding a CRF layer at bottom.



Evaluation and Comparison

Conclusion