Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

- Towards a Join Entity Linking Pipeline
- Joint Entity Linking with BERT

Evaluation

- AIDA-CoNLI Dataset
- Evaluation Criteria
- Evaluation Details

Joint Entity Linking with BERT

Amund Faller Råheim

Master of Science in Computer Science Albert-Ludwigs-Universität Freiburg

1st of June 2022

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

- Towards a Joint Entity Linking Pipeline
- Joint Entity Linking with BERT

Evaluation

- AIDA-CoNLL Dataset
- Evaluation Criteria
- Evaluation Details

1 Problem

2 Solution

3 Evaluation



◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 - のへで

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Entity Linking: Connecting an Input Document to a Knowledge Base

Problem

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

- Towards a Joint Entity Linking Pipeline
- Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Entity Linking: Connecting an Input Document to a Knowledge Base

Example

• Paris Hilton is visiting Paris this weekend.

Problem

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

- Towards a Joint Entity Linking Pipeline BERT
- Joint Entity Linking with BERT

Evaluation

- AIDA-CoNLI Dataset
- Evaluation Criteria
- Evaluation Details

Entity Linking: Connecting an Input Document to a Knowledge Base

Example

• Paris Hilton is visiting Paris this weekend.





Problem



Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

- Towards a Joint Entity Linking Pipeline BERT
- Joint Entity Linking with BERT

Evaluation

- AIDA-CoNLL Dataset
- Evaluation Criteria
- Evaluation Details

Entity Linking: Connecting an Input Document to a Knowledge Base

Example

• Paris Hilton is visiting Paris this weekend.





Entities in the Knowledge Base

Problem

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline BERT Joint Entity Lin

with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Problem

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

${\sf Entity}\ {\sf Linking} = {\sf Mention}\ {\sf Detection} + {\sf Entity}\ {\sf Disambiguation}$



Amund Faller Råheim

The Classic Pineline Networks

The Classic Pipeline

- Mention Detection
- 2 Candidate Generation
- \leftarrow Document
- \leftarrow Mention Texts

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joir Entity Linking Pipeline

Joint Entity Linki with BERT

Evaluation

AIDA-CoNL Dataset

Evaluation Criteria

Evaluation Details

Mention Detection

- 2 Candidate Generation
- $\leftarrow \mathsf{Document}$
- $\leftarrow \mathsf{Mention} \; \mathsf{Texts}$

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

... three disjoint methods! 😫

The Classic Pipeline

Amund Faller Råheim

Problem

The Classic Pipeline

Artificial Neural Networks

Problem Summar

Solution

Towards a Join Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Artificial Neural Networks

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ のQ@

SOTA for Entity Linking: Deep Learning!

Amund Faller Råheim

Problem

The Classic Pipeline

Artificial Neural Networks

Problem Summar

Solution

Towards a Join Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNL Dataset

Evaluation Criteria

Evaluation Details

Artificial Neural Networks

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

SOTA for Entity Linking: Deep Learning!

1 ANN for Mention Detection

- 2 Mapping for Candidate Generation
- 3 ANN for Entity Disambiguation

Amund Faller Råheim

Artificial Neural Networks

Artificial Neural Networks

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ のQで

SOTA for Entity Linking: Deep Learning!

- ANN for Mention Detection ⊼ 💥 2 Mapping for Candidate Generation
- **3** ANN for Entity Disambiguation ↑ Gradient

No end-to-end deep learning with a mapping in the middle!

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks

Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Questions

Problem Summary

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ 三三 - のへぐ

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks

Problem Summary

Solution

Towards a Join Entity Linking Pipeline BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNL Dataset

Evaluation Criteria

Evaluation Details

Questions

Summary

ANN for Mention Detection, Mapping for Candidate Generation, ANN for Entity Disambiguation

No benefits of end-to-end learning!

Problem Summary

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Towards a Joint Entity Linking Pipeline

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ のQ@

Joint Mention Detection and Entity Disambiguation

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Towards a Joint Entity Linking Pipeline

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Joint Mention Detection and Entity Disambiguation

Q: How to solve both tasks with one model?

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Towards a Joint Entity Linking Pipeline

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ のQで

Joint Mention Detection and Entity Disambiguation

Q: How to solve both tasks with one model?

A: Mention Embeddings and Entity Embeddings from shared Word Embeddings

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

A Transformer Network

Gives a contextualized embedding of words

BERT

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ のQ@

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

A Transformer Network

Gives a contextualized embedding of words using "self-attention":



BERT

◆□ ▶ ◆□ ▶ ◆ □ ▶ ◆ □ ▶ ○ ○ ○ ○ ○

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

A Transformer Network

Gives a contextualized embedding of words using "self-attention":



Pre-trained for general language tasks, Fine-tuned for specific tasks

BERT

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Join Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Joint Entity Linking with BERT



▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joir Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Joint Entity Linking with BERT



・ロト・日本・日本・日本・日本・日本

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joir Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Joint Entity Linking with BERT



▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

1 Mention Detection & Entity Disambiguation

(Candidate Generation)

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLI Dataset

Evaluation Criteria

Evaluation Details

Questions?

Solution Summary

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 - のへで

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Join Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Solution Summary

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Questions?

Solution Summary

Joint representation from BERT, with a Mention Detection prediction and an Entity Embedding prediction.

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Join Entity Linking Pipeline

DERI

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Questions?

Solution Summary

Joint representation from BERT, with a Mention Detection prediction and an Entity Embedding prediction.

Entity Embeddings compared with Wikipedia2vec target entity embedding.

Solution Summary

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Amund Faller Råheim

Demo

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joir Entity Linking Pipeline

BERT

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

AIDA-CoNLL Dataset

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 - のへで

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

- Towards a Joint Entity Linking Pipeline BERT
- Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

AIDA-CoNLL Dataset

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ のQで

Demo

HAVEL PRAISES CZECH NATIVE ALBRIGHT AS FRIEND . Klara Gajduskova PRAGUE 1996-12-06 Czech President Vaclav Havel on Friday welcomed the appointment of Madeleine Albright, who is of Czech extraction, as the United States ' first woman Secretary of State . In a statement Havel, who is recovering from cancer surgery, said : "Madeleine Albright is a distinguished friend, a tested diplomat, and a true American of fine origins . " " I look forward to continuing our good relations ... with the United States and with the first woman ever to hold the position of Secretary of State . I wish her well , " Have said in a statement to Reuters . Have , who helped lead the " velvet revolution " that ousted the Communist regime in Prague in 1989, invited Albright, then working for a private foreign policy think tank, to advise his new democratic government in 1990. Havel had a small malignant tumour removed from his lung on Monday and is recovering in hospital. Albright, born Marie Korbelova to a Czechoslovak diplomat in 1937, fled with her family to the United States after the Communists came to power in a coup in 1948. As an academic, Albright studied and lectured on Europe 's 20th century problems before becoming U.S. ambassador to the United Nations. Czech diplomats, seeking to have their country included in the expected expansion of NATO, praised the selection of Albright, known to be a strong supporter of alliance 's integration of former Soveit-bloc countries . " The nomination ... is a clear signal that one key of the lines of foreign policy will be the strengthening of the trans-Atlantic cooperation, a creation of strategic partnership between Europe and the US, "Foreign Minister Josef Zieleniec told Reuters, " (Albright) is a convinced advocate of NATO enlargement and of stabilisation of security structures, " Czech ambassador to the United Nations, Karel Kovanda, told the daily Mlada Fronta Dnes that Albright " is a little light in our diplomatic heaven. " but warned against expecting her to exert any influence in favour of the Czechs .

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Join Entity Linking Pipeline

Joint Entity Linkin with BERT

Evaluation

AIDA-CoNLI Dataset

Evaluation Criteria

Evaluation Details

Evaluation Criteria

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

• Precision:

 $\mathsf{Precision} = \frac{\mathsf{True} \; \mathsf{Positive}}{\mathsf{True} \; \mathsf{Positive} + \mathsf{False} \; \mathsf{Positive}}$

Recall:

$$\mathsf{Recall} = \frac{\mathsf{True \ Positive}}{\mathsf{True \ Positive} + \mathsf{False \ Negative}}$$

• F1 score:

F1 score =
$$2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Join Entity Linking Pipeline

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLI Dataset

Evaluation Criteria

Evaluation Details

Evaluation Criteria

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

• Precision:

 $\mathsf{Precision} = \frac{\mathsf{True} \; \mathsf{Positive}}{\mathsf{True} \; \mathsf{Positive} + \mathsf{False} \; \mathsf{Positive}}$

• Recall:

$$\mathsf{Recall} = \frac{\mathsf{True} \; \mathsf{Positive}}{\mathsf{True} \; \mathsf{Positive} + \mathsf{False} \; \mathsf{Negative}}$$

• F1 score:

F1 score =
$$2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

In-KB F1 Score: Ignores unknown entities

F

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Join Entity Linking Pipeline

Jaine East

with BERT

Evaluation

AIDA-CoNLI Dataset

Evaluation Criteria

Evaluation Details

Evaluation Results

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ のQ@

Results without Candidate Generation:

	AIDA-CoNLL Test (Micro F1 Score)	
	Mention Detection	Entity Linking
Our Model	95.1	56.4
Chen et al. (2019)		69.4
Broscheit (2019)		79.3

Evaluation Results

Joint Entity Linking with BERT

Amund Faller Råheim

Problem

The Classic Pipeline
Artificial Neural
Networks
Problem Summary

Solution

Towards a Join Entity Linking Pipeline

Joint Entity Linkir with BERT

Evaluation

AIDA-CoNLI Dataset

Evaluation Criteria

Evaluation Details

Results with Candidate Generation:

	AIDA-CoNLL Test (Micro F1 Score)	
	Mention Detection	Entity Linking
Broscheit (2019)		79.3
Martins et al. (2019)	92.5	81.9
Kolitsas et al. (2018)		82.4
Our Model	95.1	83.0
Poerner et al. (2020)		85.0
Chen et al. (2019)		87.7

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline BERT Joint Entity Linki with BERT

AIDA-CoNLI Dataset

Evaluation Criteria

Evaluation Details

Evaluation Details

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ のQ@

Performance on entities *seen* in the training set vs. *unseen* in the training set:

	Entity Disambiguation Accuracy (%)		
	All Mentions	Seen	Unseen
Our Model - No CG	59.4	93.1	7.6
Our Model - CG	88.1	97.2	74.6

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Evaluation Summary

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ のQ@

Evaluation Summary

• Promising results on *seen* entities without Candidate Generation,

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLI Dataset

Evaluation Criteria

Evaluation Details

Evaluation Summary

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Evaluation Summary

- Promising results on *seen* entities without Candidate Generation,
- Large performance boost with Candidate Generation.

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLI Dataset

Evaluation Criteria

Evaluation Details

Evaluation Summary

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Evaluation Summary

- Promising results on *seen* entities without Candidate Generation,
- Large performance boost with Candidate Generation.

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline

Joint Entity Linking with BERT

Evaluation

AIDA-CoNLI Dataset

Evaluation Criteria

Evaluation Details

Evaluation Summary

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Evaluation Summary

- Promising results on *seen* entities without Candidate Generation,
- Large performance boost with Candidate Generation.

Questions?

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Join Entity Linking Pipeline BERT

Joint Entity I

Evaluation

AIDA-CoNLI Dataset

Evaluation Criteria

Evaluation Details

◆□ ▶ ◆□ ▶ ◆ 臣 ▶ ◆ 臣 ▶ ○ 臣 ○ のへで

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline BERT Joint Entity Link with BERT

Evaluation

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

AIDA-CoNLL Stats (Table 2)

Dataset	Docs	Ment	Ment Annot	Unique Ent
Train	946	23,396	18,541	4,084
Validation	216	5,917	4,791	1,644
Test	231	5,616	4,485	1,536
Total	1,393	34,929	27,817	5,593

Table: Number of documents, mentions, mentions annotated with Wikipedia entities and unique mentioned entities in the AIDA-CoNLL datasets.

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Amund Faller Råheim

Problem

The Classic Pipeline Artificial Neural Networks Problem Summary

Solution

Towards a Joint Entity Linking Pipeline BERT Joint Entity Linki with BERT

AIDA-CoNLL Dataset

Evaluation Criteria

Evaluation Details

Model Comparisons (Table 12)

	Our Model	Chen et al. (2019)
BERT Hidden Lavers	Uncased Yes	Cased No
Training Epochs	180	190
Training Time	4 hrs 50 mins	5 hrs 6 mins
Loss Function λ	0.01	0.1
Dropout	No	Yes

Table: Characteristics of our best-performing model and our implementation with the settings of Chen et al. (2019).