DeNERT-KG
논문 리뷰

Related Work

- **NER : Named Entity Recognition**

1) Rule-based : Based on rules manually defined by people
   - Irregular and Incomplete : because of nature of natural language
   - Highly likely to work well only in certain datasets

2) Dictionary-based : Based on collected dictionaries or user-defined dictionary
   - Advantage : Information extraction or retrieval in certain area
   - Disadvantage : Have to organize, costly to manage,
     limitation to non-pre-defined-words

=> Utilize DQN & BERT : improve the performance of NER task model

Related Work

- **Relation Extraction**

Extract the relationship between two words within sentences and documents

- **Joint Task of Extracting Entities and Relationships**

Joint task can be performed depending on NLP tools

1. RNN : Encode linguistic and syntactic properties in text
2. CNN : Better capture semantic information in sentences

Limitation : Focus only on the relationship between the entities

=> Cannot capture information about the types of entities in sentence

Related Work

- **DQN : Deep Q - Networks**

  Neural Networks + Reinforcement Learning

  => Solve the problem of training
  not being converged well
  by approximating Q value
  using neural networks

- **BERT**

  Pre-trained Language model

  1) Train language expressions
     through unsupervised learning

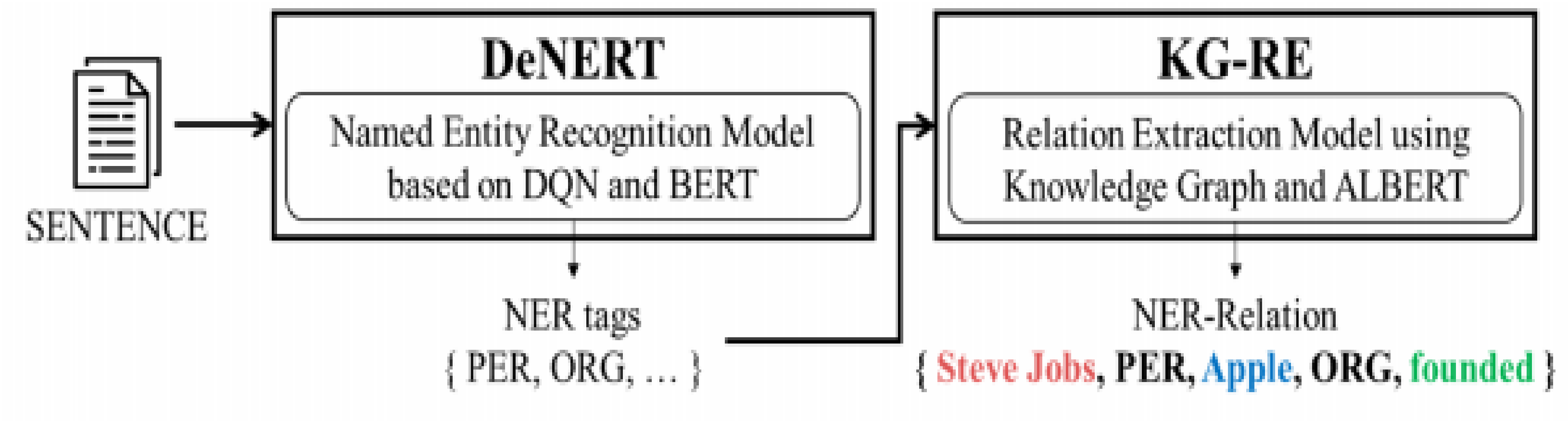
  2) Train model for specific downstream tasks

- **Knowledge Graph**

  Graph of words linking them together and can help computer learn person’s common-sense more easily

DeNERT - KG

- **Architechture**

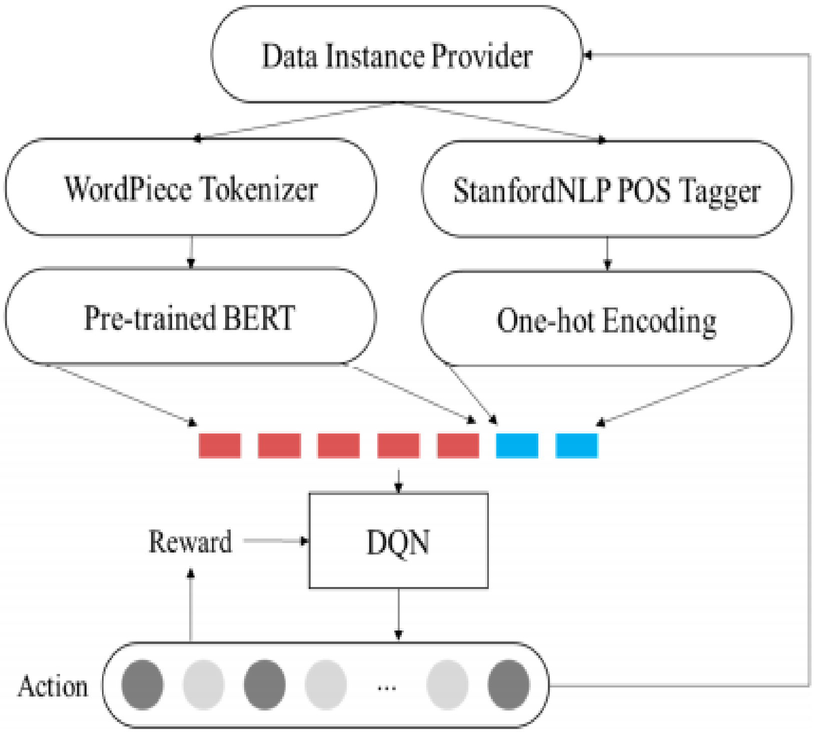


DeNERT : Proceed NER based on DQN & BERT

KG - RE : Proceed Relation Extraction based on Knowledge Graph and ALBERT

DeNERT - KG

- **DeNERT**



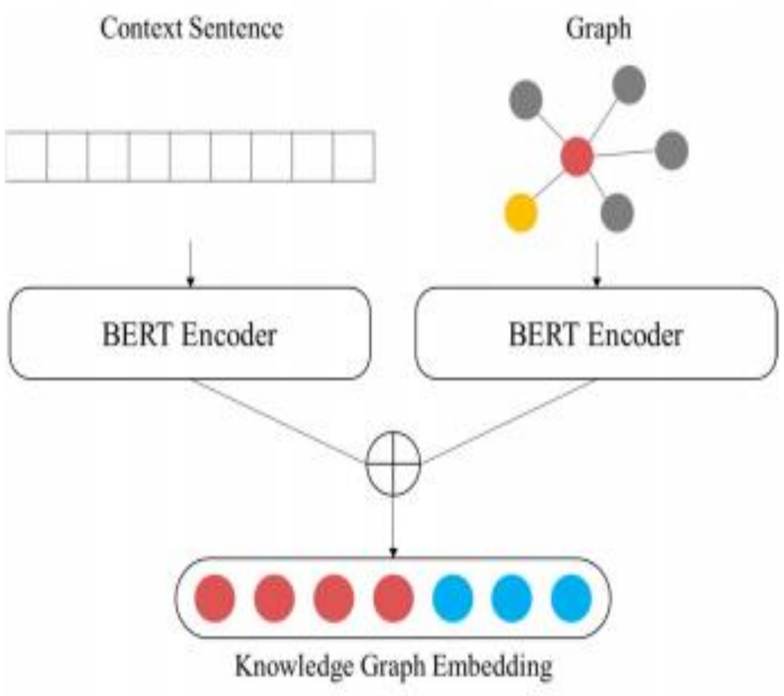
Apply the reinforcement learning model
to classification task

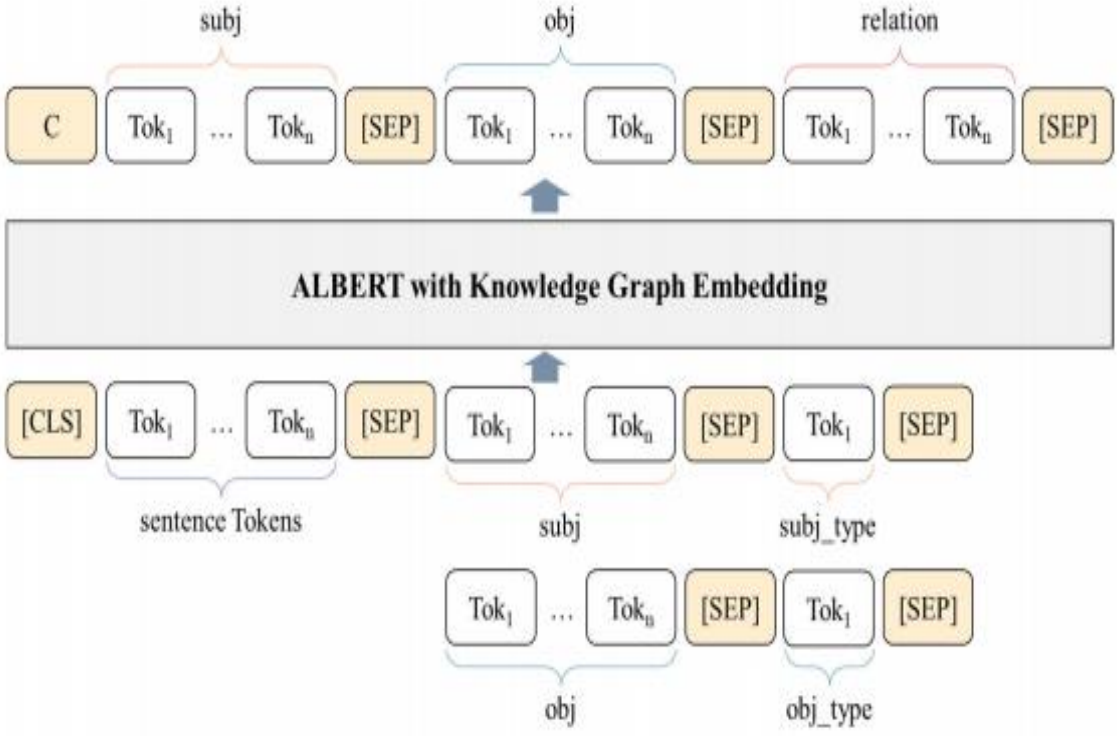
=> Important elements in reinforcement learning agent
should be redefined to suit the NER problem

: **State, Action, Reward**

DeNERT - KG

- **Knowledge-Graph based Relation Extraction Model**





Experiments

- **Result for NER Task**

- DQN by conforming BERT & POS

Table 4. Experiment result on EWNERTC dataset.

Model	Precision	Recall	F1
BERT-Base	77.52	70.58	73.89
BERT-Base + POS	78.82	70.84	74.62
BERT-Large	80.25	73.89	76.94
BERT-Large + POS (DeNERT)	81.97	74.13	**77.85**

DeNERT model shows
The highest F1-score

- Comparative experiment with existing models

Table 5. Comparison result on CoNLL 2003 and W-NUT 2017 dataset.

Dataset	Model	F1-Score
CoNLL 2003	Akbik et al., 2019 [38]	93.18
	Brian et al., 2020 [39]	91.47
	Peters et al., 2017 [40]	91.93
	DeNERT	**93.45**
W-NUT 2017	Pius et al., 2017 [41]	41.06
	Gustavo et al., 2019 [37]	41.86
	Akbik et al., 2019 [38]	**49.59**
	Brian et al., 2020 [39]	40.59
	DeNERT	45.78

1) CoNLL : highest F1-score
2) W-NUT : relatively lower results
   because of many noise data

Experiments

- **Result for Relation Extraction Task**

- Verify performance
  When using Knowledge Graph

Table 6. Experiment result on TACRED dataset.

Model	Precision	Recall	F1-Score
BERT-only	66.4	66.8	66.6
BERT + KG	72.6	68.3	70.4
Electra-base + KG	48.2	43.6	45.8
ALBERT + KG (DeNERT-KG)	**71.8**	**73.1**	**72.4**

Best results when we used
the DeNERT-KG and ALBERT

- Comparative experiment with existing models

Table 7. Comparison result on TACRED dataset.

Model	Precision	Recall	F1-Score
Matching-the-Blanks [23]	-	-	71.5
C-GCN + PA-LSTM [22]	71.3	65.4	68.2
PA-LSTM [21]	65.7	64.5	65.1
Logistic Regression [22]	73.5	49.9	59.4
ALBERT + KG (DeNERT-KG)	**71.8**	**73.1**	**72.4**

Propose model didn't achieve higher
performance than existing one

But, existing model extract only relevant
information from sentences

THANK YOU