FUJITSU
shaping tomorrow with you

UOM-FJ at the NTCIR-15 SHINRA2020-ML Task

Hiyori Yoshikawa [1] Chunpeng Ma[1] Aili Shen [2] Qian Sun [2]
Chenbang Huang [2] Guillaume Pelat [3] Akiba Miura [1] Daniel Beck [2]
Timothy Baldwin [2] Tomoya Iwakura [1]

[1]Fujitsu Laboratories Ltd., Japan [2]University of Melbourne, Australia
[3]École Polytechnique, France

Summary

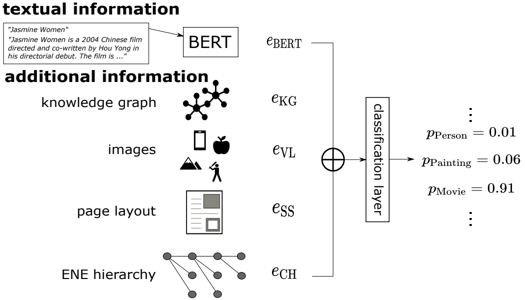
- We participated in 28 subtasks of the **SHINRA2020-ML Task**: mapping Wikipedia entities into Extended Named Entity (ENE) categories.
- Our model was trained to capture multiple aspects of Wikipedia articles: **text, structured knowledge, images, page layout and the ENE class hierarchy**.
- Our system **ranked first in four languages**, achieving an F1-score of 82.73 on the English subtask.

Our System Overview

Document representation: $e = \mathrm{FFNN}(e_{\mathrm{BERT}} \oplus e_{\mathrm{KG}} \oplus e_{\mathrm{VL}} \oplus e_{\mathrm{SS}} \oplus e_{\mathrm{CH}})$

- BERT encoder (Devlin et al., 2019) for **pure textual information** (e_{BERT})
- A set of separately-trained document representations that encodes different aspects of a given document:
  knowledge graph features (e_{KG}), text and images (e_{VL}), page screenshot layout (e_{SS}), and ENE class hierarchy (e_{CH})

Label probability: $p(c, e) = \sigma(w_c^{\top} e + b_c)$

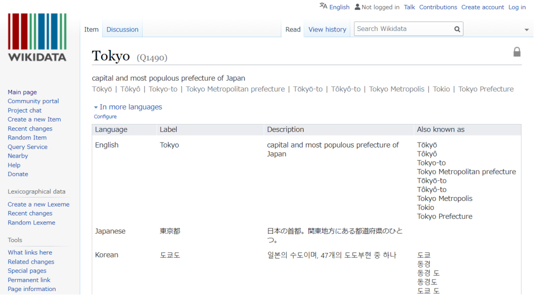


Multi-aspect Document Representations (1/4)

FUJITSU

Knowledge graph (e_{KG})

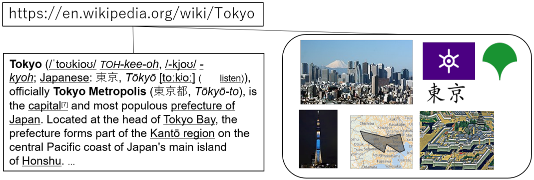
- Pre-trained embeddings of **Wikidata graph** (Lerer et al., 2019)
- `wikibase_item` field of the Wikipedia dump is used to identify the Wikidata entity corresponding to a Wikipedia article
- 98.3% of English Wikipedia articles have corresponding Wikidata entries



`https://www.wikidata.org/wiki/Q1490`

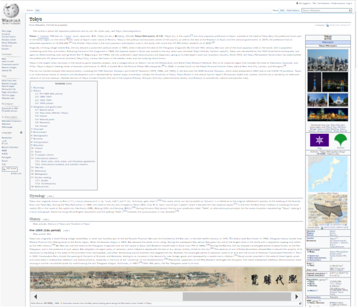
Text–image representation (e_{VL})

- **VL-BERT**-based (Su et al., 2020) encoding of **text and images in a given Wikipedia article**
- Multiple images are rescaled and concatenated to compose a single large image; individual images are treated as regions-of-interest (ROIs)
- The VL-BERT model is fine-tuned on the SHINRA2020-ML task



Page screenshot layout (e_{SS})

- We obtain **visual renderings of Wikipedia articles** from a Wikipedia dump to generate their screenshots
- The screenshots are resized to a fixed size and fed into INCEPTION (Inception V3 with different convolution filters) (Szegedy et al., 2016) to obtain visual representations



Multi-aspect Document Representations (4/4)

FUJITSU

ENE class hierarchy (e_{CH})

- We employ a **hierarchy-aware global model (HiGAM)** (Zhou et al., 2020) to capture **the label hierarchy and correlations between ENE categories**
- Construct hierarchy graph: node = ENE category, edge = hierarchical relational information or correlation information
- The encoder is trained with binary cross-entropy loss over the hierarchical label space

http://ene-project.info/ene8/?lang=en

Submissions

We examined three ways to generate the final outputs:

- `jointrep`: output the categories whose estimated probability exceeds the threshold $\theta = 0.5$
- `jointrepPostprocess`: `jointrep` + assign the label `CONCEPT` if the probability scores of all candidate categories are below the threshold
- `jointrepUnionPostprocess`: take the union of different predictions made by five models trained on different subsets of the training data, and then apply the post-processing step as in `jointrepPostprocess`

Main Results

FUJITSU

Official evaluation results (English)

Submission name	$\mathcal{P}$	$\mathcal{R}$	$\mathcal{F}_1$	Rank
`jointrep`	**81.77**	83.71	**82.73**	1
`jointrepPostprocess`	81.46	83.71	82.57	3
`jointrepUnionPostprocess`	80.66	84.80	82.68	2
Other teams best	79.65	**85.00**	82.23	

Other languages

- Non-English results are based on VL-BERT model, i.e. using only text–image representation
- Ranked first on **Spanish, Italy and Catalan** as well
- Relatively low performance on Arabic, Hindi and Thai

Ablation Study

leaderboard results, `jointrep`, English

Model	$\mathcal{P}$	$\mathcal{R}$	$\mathcal{F}_1$
Full	75.3	76.4	75.7
– knowledge graph	74.8	75.8	75.2
– text–image	74.5	75.5	74.8
– page layout	75.5	77.2	76.0
– class hierarchy	74.4	76.4	75.1
BERT + knowledge graph	70.8	71.3	71.0
BERT + text–image	73.5	74.8	73.9
BERT + page layout	70.5	71.6	70.9
BERT + class hierarchy	74.2	75.2	74.5

Error examples

FUJITSU

Page ID	Title	Opening text	Predicted ENEs	Gold ENEs
54866853	Hedi (Policy)	Hedi was an economic policy of the imperial China. It was the state purchase of food supplies from farmers. As a means to control the price of grains and foods, it is an early example of Government procurement. The policy was adopted in the year of 488 by Emperor Xiaowen of Northern Wei as a counter measure of drought. The state purchases food supplies and stock them. ...	1.7.21.8:Plan	0:CONCEPT
299259	Roman currency	Roman currency for most of Roman history consisted of gold, silver, bronze, orichalcum and copper coinage (see: Roman metallurgy). From its introduction to the Republic, during t Imperial times, Roman currency saw m ination, and composition. A persistent debasement and replacement of coins over the centuries. ...	1.7.25.1:Currency	9:IGNORED
788538	Azamino	Azamino (あざみ野) is a bedroom community of Tokyo and Yokohama, located in Aoba-ku, Yokohama, Kanagawa Prefecture, Japan. The area is 20 minutes from Shibuya Station on the Tōkyū Den-en-toshi Line and 27 minutes by Known as a refer to the years. It is located next to the trendy Tama Plaza.	1.5.0:Location_-Other	1.5.1.1:City
22786903	Hiro H1H	The Hiro H1H (or Navy Type 15) was a 1920s Japanese bomber or reconnaissance biplane flying boat developed from the Felixstowe F.5 by the Hiro Naval Arsenal for the Imperial Japanese Navy. The aircraft were built by Hiro, the Yokosuka Naval Arsenal and Aichi.	1.7.6:Weapon, 1.7.17.3:Aircraft	1.7.17.3:Aircraft
586540	Tailor	A tailor is a person who makes, repairs, or alters clothing professionally, especially suits and men's clothing. Although the term dates to the thirte nth century, and now refers to makers of men's and women's suits, coats, trousers, and similar garments, usually of wool, linen, or silk. ...	1.7.23.1:Position_-Vocation	0:CONCEPT

Confusion between CONCEPT/IGNORED and other classes

Information mismatch between target and Japanese Wikipedia articles

Redirect fromTailoring

Page redirect to different entity

↓ Tailoring

Summary (recap)

FUJITSU

- We participated in 28 subtasks of the **SHINRA2020-ML Task**: mapping Wikipedia entities into Extended Named Entity (ENE) categories.
- Our model was trained to capture multiple aspects of Wikipedia articles: **text, structured knowledge, images, page layout and the ENE class hierarchy**.
- Our system **ranked first in four languages**, achieving an F1-score of 82.73 on the English subtask.

References

J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, June 2019. doi: 10.18653/v1/N19-1423. URL https://www.aclweb.org/anthology/N19-1423.

A. Lerer, L. Wu, J. Shen, T. Lacroix, L. Wehrstedt, A. Bose, and A. Peysakhovich. PyTorch-BigGraph: A large-scale graph embedding system. In *Proceedings of the 2nd SysML Conference*, 2019.

W. Su, X. Zhu, Y. Cao, B. Li, L. Lu, F. Wei, and J. Dai. Vl-bert: Pre-training of generic visual-linguistic representations. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=SygXPaEYvH.

C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2818–2826, 2016.

J. Zhou, C. Ma, D. Long, G. Xu, N. Ding, H. Zhang, P. Xie, and G. Liu. Hierarchy-aware global model for hierarchical text classification. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 1106–1117, 2020.

shaping tomorrow with you