**Summary**

- We participated in 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 = \text{FFNN}(e_{\text{BERT}} \oplus e_{\text{KG}} \oplus e_{\text{VL}} \oplus e_{\text{SS}} \oplus e_{\text{CH}})$

- BERT encoder [1] for pure textual information ($e_{\text{BERT}}$)
- A set of separately-trained document representations that encodes different aspects of a given document: knowledge graph features ($e_{\text{KG}}$), text and images ($e_{\text{VL}}$), page screenshot layout ($e_{\text{SS}}$), and ENE class hierarchy ($e_{\text{CH}}$)

**Label probability:** $p(c, e) = \sigma(w_c^Te + b_c)$

- BERT encoder is fine-tuned while training the classifier
- Other document representations are separately trained, and fixed during training of the final classifier

**Knowledge graph ($e_{\text{KG}}$)**

- We use pre-trained embeddings of a Wikidata graph [3]
- The 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

**Text–image representation ($e_{\text{VL}}$)**

- 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 and used to generate the text–image representation

**Page screenshot layout ($e_{\text{SS}}$)**

- We obtain visual renderings of Wikipedia articles from a Wikipedia dump and generate their screenshots
- The screenshots are resized to a fixed size and fed into *INCEPTION* (Inception V3 with different convolution filters) [5] to obtain visual representations

**ENE class hierarchy ($e_{\text{CH}}$)**

- We employ a hierarchy-aware global model (HiGAM) [6] to capture the label hierarchy and correlations between ENE categories
- Textual representation of a document is fed into a Hierarchy-GCN [2], where each node of the hierarchy graph represents an ENE category, and each directed edge represents either hierarchical relational information or correlation information, to obtain hierarchy-aware embeddings
- The encoder is trained with binary cross-entropy loss over the hierarchical label space

**Results**

### Official evaluation results (English)

<table>
<thead>
<tr>
<th>Submission name</th>
<th>F1</th>
<th>R</th>
<th>F1 Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>jointrep</td>
<td>81.77</td>
<td>83.71</td>
<td>82.73</td>
</tr>
<tr>
<td>jointrepPostprocess</td>
<td>81.46</td>
<td>83.71</td>
<td>82.57</td>
</tr>
<tr>
<td>jointrepUnionPostprocess</td>
<td>80.66</td>
<td>84.80</td>
<td>82.68</td>
</tr>
<tr>
<td>Other teams best</td>
<td>79.65</td>
<td>85.00</td>
<td>82.23</td>
</tr>
</tbody>
</table>

### Official evaluation results (Other languages*, F1)

<table>
<thead>
<tr>
<th>Language</th>
<th>F1</th>
<th>R</th>
<th>F1 Rank</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>ar</td>
<td>64.55</td>
<td>83.07</td>
<td>79.62</td>
<td>81.29</td>
</tr>
<tr>
<td>Others best</td>
<td>75.27</td>
<td>83.77</td>
<td>76.28</td>
<td>84.47</td>
</tr>
</tbody>
</table>

**Analysis**

### Ablation study (leaderboard results, jointrep, English)

<table>
<thead>
<tr>
<th></th>
<th>Full</th>
<th>$e_{\text{KG}}$</th>
<th>$e_{\text{VL}}$</th>
<th>$e_{\text{SS}}$</th>
<th>$e_{\text{CH}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mathcal{F}_1$</td>
<td>75.7</td>
<td>75.2</td>
<td>74.8</td>
<td>76.0</td>
</tr>
<tr>
<td></td>
<td>$\mathcal{F}_1$</td>
<td>75.5</td>
<td>75.2</td>
<td>74.8</td>
<td>76.0</td>
</tr>
</tbody>
</table>

**Common error patterns**

- Confusion between **CONCEPT** and other classes, e.g. 1.7.21: Doctrine,Method
- Informatin mismatch: some information in the corresponding Japanese article is missing in the target article, and vice versa
- Page redirects to a different entity, e.g. Tailoring → Tailor

**References**


