Neural Information Retrieval for Educational Resources
ECIR 2022 – Industry Day

Background and Motivation

- **Wizenoze - The Company**
  - “Wizenoze brings structure in a world of information overload”
  - EdTech provider since 2013

- **Product**
  - API providing educational content
  - Content selection: human curators plus AI / NLP
    - Tailored for custom-defined curricula
    - Backed by search engine
  - **Better search results -> less human curation effort**

Train a Sentence BERT Model for Educational Resources

- Fine-tuned Sentence BERT on historic curation data
  - ~33k in-domain query-document-judgement triplets
- Model selection on held-out validation set
- Vanilla S-BERT vs fine-tuned S-BERT on test sub-set to evaluate training impact

<table>
<thead>
<tr>
<th>Model</th>
<th>P@5</th>
<th>P@10</th>
<th>MAP@5</th>
<th>MAP@10</th>
</tr>
</thead>
<tbody>
<tr>
<td>MiniLM-L6-v2</td>
<td>0.7926</td>
<td>0.7888</td>
<td>0.8505</td>
<td>0.8504</td>
</tr>
<tr>
<td>mpnet-base-v2</td>
<td>0.7851</td>
<td>0.7926</td>
<td>0.8619</td>
<td>0.8654</td>
</tr>
<tr>
<td>Fine-tuned</td>
<td><strong>0.8889</strong></td>
<td><strong>0.8667</strong></td>
<td><strong>0.9370</strong></td>
<td><strong>0.9260</strong></td>
</tr>
</tbody>
</table>

Evaluation and Validation

- Evaluate on separate test set:
  - 39 hand-picked test queries
- Sanity checks: manually validate 50 most different results sets (RBO score)
- A/B testing (WIP):
  - which model is preferred by curators?

<table>
<thead>
<tr>
<th>Model</th>
<th>P@5</th>
<th>P@10</th>
<th>MAP@5</th>
<th>MAP@10</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM25+</td>
<td>0.6963</td>
<td>0.6407</td>
<td>0.8402</td>
<td>0.8039</td>
</tr>
<tr>
<td>Fine-tuned</td>
<td><strong>0.7926</strong></td>
<td><strong>0.7519</strong></td>
<td><strong>0.8727</strong></td>
<td><strong>0.8632</strong></td>
</tr>
</tbody>
</table>

Summary and Conclusions

- Neural Information Retrieval: significant improvements on education domain search
- Learning in-domain expert preferences
- Affordable with serverless implementation, gradual transition