Hunter Gatherer: UDEM at 1CLICK-2

Pablo Duboue, Jing He, Jian-Yun Nie
DIRO, Université de Montréal, Canada

Introduction

• 1CLICK-2 Task: retrieve relevant information Units instead of relevant documents.
• Hunter Gatherer System
  ✓ DeepQA framework for vital string ranking
  ✓ ILP or MMR for summarization

Evidence Gatherer

• Input: Candidate Set
• Output: Candidate Scores
• Method:
  a) Gather evidence
     - Build Indri Query from the original query and candidate
     - Evaluate the query and get top ranked passages and scores
  a) Score candidates based on evidences

\[
R(q, u) = \lambda_1 \sum_{p \in E, a \in p} (R(q, p) + \alpha) + \lambda_2 \sum_{p \in E} (R(q, p) + \beta)
\]

Summarizer: Organizing Candidates

• Input: Candidate Scores
• Output: Constraint length of text containing most relevant candidates
• Method:
  – MMR (Maximal Marginal Relevance) iterative algorithm to select sentences
  – Score sum of candidates in a sentence as the relevant score
  – Jaccard distance of bigrams between sentences as redundancy score
  – ILP (Integer Linear Programming)
    – Objective function: score sum of candidates in the selected sentences
    – Constraint: the length of sentences is smaller than value k.

Evaluation

• Run 1. Hunter (candidate = tokens + NEs) + Gatherer + Summarizer (MMR)
• Run 3. Hunter (candidate = tokens + NEs + learnt extractor output) + Gatherer + Summarizer (MMR)
• Run 4. Hunter (candidate = tokens + NEs) + Gatherer + Summarizer (ILP)

<table>
<thead>
<tr>
<th>Run / Category</th>
<th>All</th>
<th>ACTOR</th>
<th>ATHLETE</th>
<th>ARTIST</th>
<th>POLITICIAN</th>
<th>FACULTY</th>
<th>GEO</th>
<th>DEFEND</th>
<th>QA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run 1</td>
<td>0.047</td>
<td>0.040</td>
<td>0.028</td>
<td>0.039</td>
<td>0.037</td>
<td>0.060</td>
<td>0.025</td>
<td>0.066</td>
<td>0.068</td>
</tr>
<tr>
<td>Run 3</td>
<td>0.050</td>
<td>0.058</td>
<td>0.016</td>
<td>0.038</td>
<td>0.086</td>
<td>0.058</td>
<td>0.016</td>
<td>0.077</td>
<td>0.053</td>
</tr>
<tr>
<td>Run 4</td>
<td>0.080</td>
<td>0.086</td>
<td>0.084</td>
<td>0.074</td>
<td>0.025</td>
<td>0.079</td>
<td>0.062</td>
<td>0.076</td>
<td>0.146</td>
</tr>
<tr>
<td>MAX</td>
<td>0.080</td>
<td>0.086</td>
<td>0.084</td>
<td>0.074</td>
<td>0.086</td>
<td>0.083</td>
<td>0.080</td>
<td>0.088</td>
<td>0.146</td>
</tr>
<tr>
<td>MIN</td>
<td>0.047</td>
<td>0.040</td>
<td>0.016</td>
<td>0.018</td>
<td>0.025</td>
<td>0.005</td>
<td>0.016</td>
<td>0.055</td>
<td>0.053</td>
</tr>
<tr>
<td>AVRG</td>
<td>0.059</td>
<td>0.053</td>
<td>0.034</td>
<td>0.032</td>
<td>0.049</td>
<td>0.070</td>
<td>0.044</td>
<td>0.067</td>
<td>0.096</td>
</tr>
<tr>
<td>MEDIAN</td>
<td>0.055</td>
<td>0.053</td>
<td>0.028</td>
<td>0.027</td>
<td>0.039</td>
<td>0.076</td>
<td>0.035</td>
<td>0.096</td>
<td>0.089</td>
</tr>
</tbody>
</table>

References