The higher the M score is, the more we will rely on this level of abstraction.

**Motivation**

- Users query may be of different nature
- Lexical queries: Ron Howard
- VS. Conceptual queries: Lymphoma in dogs

How?

- Combine the interactions of different levels of abstraction by a gating mechanism
- Weak supervision to generate enough training data

**Methodology**

Multi-level Abstraction Convolutional Model (MACM)

- Input interaction matrix \( I_{ij} = \cos(t_{i}^{(q)}, t_{j}^{(d)}) \)
- 2D convolutions and max-pooling layers
- Flatten each max-pooled layer and apply MLP to obtain a level-specific matching score \( S_{i} \)
- Level-importance feature M:
  \[
  M^{(i)}(k) = \max_{v=1..m}[C^{(k)}]_{uv} \quad M^{(i)} = \sum_{u=1}^{n} M^{(i)}(k)
  \]
  The higher the M score is, the more we will rely on this level of abstraction.
- Softmax normalization gate:
  \[
  \beta_{k} = \frac{e^{\exp(\alpha_{k} M^{(i)})}}{\exp(\sum_{j=0} M^{(j)})}
  \]
- Final score prediction: by MLP \( S = f(W[\beta_{0}S_{0}, ..., \beta_{L}S_{L}] + b) \)

**Experiments**

<table>
<thead>
<tr>
<th>Collection</th>
<th>Genre</th>
<th>Validation Queries</th>
<th>Test Queries</th>
<th>#Docs</th>
<th>Avg.d.Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clueweb09B</td>
<td>Webpages</td>
<td>1-50</td>
<td>51-200</td>
<td>50M</td>
<td>1,506</td>
</tr>
</tbody>
</table>

- Previous models only employ the representation/interaction score at the same level of abstraction for any query.
- Users queries may be of different nature: lexical queries, conceptual queries, and queries in between
- We propose MACM which employs interaction scores at multiple levels of abstraction to estimate relevance
- Experimental results show that MACM outperforms the BM25 baseline and demonstrate the effectiveness of multi-level matching over single-level matching.