**Task 1: Evaluation Metrics and α-NDCG**

1.1 Name two aspects that are not considered when using the standard information retrieval evaluation metrics precision and recall to measure the relevance of a ranked document by a search engine.

- Ranking
- Diversity

1.2 Given the equations for α-NDCG@k and for the gain (for explanations see appendix), answer the following questions for each letter in “α-NDCG”:

- What aspect does the letter stand for?
- What is that aspect’s role in α-NDCG?
- In which part of the equations is that aspect covered?

\[
\alpha - NDCG(q, k) = Z_{kj} \sum_{m=1}^{k} \frac{G[m]}{\log_2(1 + m)}, \quad G[m] = \sum_{i=1}^{|S|} J(d_{m, n_i})(1 - \alpha)^{r_{i,m-1}}
\]

<table>
<thead>
<tr>
<th>Letter</th>
<th>Name</th>
<th>Description</th>
<th>Part of the Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>Gain</td>
<td>The information gained from a specific document at position ( m ).</td>
<td>( G[m] )</td>
</tr>
<tr>
<td>α</td>
<td>Diversification</td>
<td>When computing the gain, redundant results are discouraged.</td>
<td>((1 - \alpha)^{r_{i,k-1}})</td>
</tr>
<tr>
<td>D</td>
<td>Discounted</td>
<td>The gain is reduced logarithmically proportional to the position of the result. That means, relevant documents at the top of the list are valued more.</td>
<td>( \log_2(1 + m) )</td>
</tr>
<tr>
<td>C</td>
<td>Cumulative</td>
<td>The gain of each document at positions 1..k is added to the score.</td>
<td>( \sum_{m=1}^{k} )</td>
</tr>
<tr>
<td>N</td>
<td>Normalized</td>
<td>The score is normalized by the best possible ranking’s score.</td>
<td>( Z_{k,j} )</td>
</tr>
</tbody>
</table>
Task 2: Computing the Gain Vector for $\alpha$-NDCG

Given is the following ranking of documents and the information nuggets they contain.

<table>
<thead>
<tr>
<th>ID</th>
<th>Rank</th>
<th>Title</th>
<th>Information Nuggets</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>1</td>
<td>Carnival Re-Enters Norway Bidding</td>
<td>X</td>
<td>2</td>
</tr>
<tr>
<td>b</td>
<td>2</td>
<td>NORWEGIAN CRUISE LINE SAYS OUTLOOK IS GOOD</td>
<td>X</td>
<td>0.5</td>
</tr>
<tr>
<td>c</td>
<td>3</td>
<td>Carnival, Star Increase NCL Stake</td>
<td>X</td>
<td>0.25</td>
</tr>
<tr>
<td>d</td>
<td>4</td>
<td>Carnival, Star Solidify Control</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>e</td>
<td>5</td>
<td>HOUSTON CRUISE INDUSTRY GETS BOOST WITH...</td>
<td>X</td>
<td>2</td>
</tr>
<tr>
<td>f</td>
<td>6</td>
<td>TRAVELERS WIN IN CRUISE TUG-OF-WAR</td>
<td>X</td>
<td>0.5</td>
</tr>
<tr>
<td>g</td>
<td>7</td>
<td>ARMCHAIR QUARTERBACKS NEED... THIS CRUISE</td>
<td>X</td>
<td>1</td>
</tr>
<tr>
<td>h</td>
<td>8</td>
<td>EUROPE, CHRISTMAS ON SALE</td>
<td>X</td>
<td>0.25</td>
</tr>
<tr>
<td>i</td>
<td>9</td>
<td>TRAVEL DEALS AND DISCOUNTS</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>j</td>
<td>10</td>
<td>HAVE IT YOUR WAY ON THIS SHIP</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

2.1 Compute the gain for each document.

\[
G[1] = 0 \cdot (1 - 0.5)^0 + 1 \cdot (1 - 0.5)^0 + 0 \cdot (1 - 0.5)^0 + 1 \cdot (1 - 0.5)^0 \\
+ 0 \cdot (1 - 0.5)^0 + 0 \cdot (1 - 0.5)^0 = 2
\]

\[
G[2] = 0 \cdot (1 - 0.5)^0 + 1 \cdot (1 - 0.5)^1 + 0 \cdot (1 - 0.5)^0 + 0 \cdot (1 - 0.5)^1 \\
+ 0 \cdot (1 - 0.5)^0 + 0 \cdot (1 - 0.5)^0 = \frac{1}{2} = 0.5
\]

\[
G[3] = 0 \cdot (1 - 0.5)^0 + 1 \cdot (1 - 0.5)^2 + 0 \cdot (1 - 0.5)^0 + 0 \cdot (1 - 0.5)^1 \\
+ 0 \cdot (1 - 0.5)^0 + 0 \cdot (1 - 0.5)^0 = \frac{1}{4} = 0.25
\]

\[
G[4] = 0
\]

\[
G[5] = 1 \cdot (1 - 0.5)^0 + 1 \cdot (1 - 0.5)^0 = 2
\]

\[
G[6] = 1 \cdot (1 - 0.5)^1 = 0.5
\]

\[
G[7] = 1 \cdot (1 - 0.5)^0 = 1
\]

\[
G[8] = 1 \cdot (1 - 0.5)^2 = 0.25
\]

\[
G[9] = 0
\]

\[
G[10] = 0
\]
**Task 3: Computing $\alpha$-NDCG**

Compute the $\alpha$-NDCG for the top-10 results ($k = 10$) for a single query $q$ that returns the document ranking from Task 2.

3.1 Create the perfect ranking for that document ranking and compute the corresponding gains and its DCG for $\alpha = 0.5$.

<table>
<thead>
<tr>
<th>Document</th>
<th>Information Nuggets</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>e</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>g</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>j</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
DCG(Q, k) = \frac{2}{\log_2(1+1)} + \frac{2}{\log_2(1+2)} + \frac{1}{\log_2(1+3)} + \frac{0.5}{\log_2(1+4)} + \frac{0.5}{\log_2(1+5)} + \frac{0.25}{\log_2(1+6)} \\
+ \frac{0.25}{\log_2(1+7)} + \frac{0}{\log_2(1+8)} + \frac{0}{\log_2(1+9)} + \frac{0}{\log_2(1+10)} \\
\approx 4.343
\]

3.2 Compute the $\alpha$-NDCG of the ranking in Task 2.

\[
\alpha - NDCG(Q, k) = Z_{10,q} \cdot \sum_{m=1}^{10} \frac{G[m]}{\log_2(1+m)} \\
= \frac{1}{4.343} \cdot \left( \frac{2}{\log_2(1+1)} + \frac{0.5}{\log_2(1+2)} + \frac{0.25}{\log_2(1+3)} + \frac{0}{\log_2(1+4)} + \frac{2}{\log_2(1+5)} + \frac{0.5}{\log_2(1+6)} \\
+ \frac{1}{\log_2(1+7)} + \frac{0.25}{\log_2(1+8)} + \frac{0}{\log_2(1+9)} + \frac{0}{\log_2(1+10)} \right) \\
\approx \frac{1}{4.343} \cdot 3.804 \approx 0.858
\]
4. Diversification algorithm

Given is the following list \( L[l] \) of top-k queries ranked by relevance:

| Rank | ID   | \( P(Q|K) \) | Query Interpretation                                      |
|------|------|--------------|----------------------------------------------------------|
| 1    | \( Q_1 \) | 0.95         | \{"consideration": movie.title, "christopher guest":director.name\} |
| 2    | \( Q_2 \) | 0.87         | \{"christopher guest":director.name\}                   |
| 3    | \( Q_3 \) | 0.78         | \{"guest": movie.title, "consideration": movie.title\}   |
| 4    | \( Q_4 \) | 0.23         | \{"consideration": movie.title\}                       |

Use the diversification algorithm to return the list \( R[r] \) of the relevant and diverse query interpretations.

4.1 Use the Jaccard similarity to compute the similarity between each pair of query interpretations.
4.2 Given these queries, interpretation probabilities and similarity values, perform the diversification algorithm to generate a diversified ranking for $r = 3$ and $a = 0.5$.

$R = \langle Q_1 \rangle$

1. $\text{best\_score} = 0$
   a. $\text{Score}(Q_2) = \frac{1}{2} \cdot P(Q_2, K) - \left(1 - \frac{1}{2}\right) \sum_{q \in Q_1} \frac{\text{Sim}(Q_2, q)}{|Q_1|}$
      \[= \frac{1}{2} \cdot 0.87 - \frac{1}{2} \cdot \frac{\text{Sim}(Q_2, Q_1)}{2} = 0.435 - \frac{1}{2} \cdot \frac{1}{2} = 0.185\]
      $\implies \text{best\_score} = 0.185$
   b. $\text{Score}(Q_3) = \frac{1}{2} \cdot 0.78 - \frac{1}{2} \cdot \frac{1}{3} = 0.22\overline{3}$
      $\implies \text{best\_score} = 0.22\overline{3}$
   c. $\text{Score}(Q_4) = \frac{1}{2} \cdot 0.23 - \frac{1}{2} \cdot \frac{1}{2} = -0.135$

2. $R = \langle Q_1, Q_3 \rangle$

3. $\text{best\_score} = 0$
   a. $\text{Score}(Q_2) = \frac{1}{2} \cdot P(Q_2, K) - \left(1 - \frac{1}{2}\right) \sum_{q \in Q_1} \frac{\text{Sim}(Q_2, q)}{|Q_1|}$
      \[= \frac{1}{2} \cdot 0.87 - \frac{1}{2} \cdot \left(\frac{\text{Sim}(Q_2, Q_1)}{2} + \frac{\text{Sim}(Q_2, Q_4)}{2}\right)\]
      \[= 0.435 - \frac{1}{2} \cdot \left(\frac{1}{2} + \frac{1}{2}\right) = 0.31\]
      $\implies \text{best\_score} = 0.31$
   b. $\text{Score}(Q_4) \leq \frac{1}{2} \cdot 0.23 = 0.115 \leq 0.31 = \text{best\_score}$
      $\implies \text{break}$
   c. $R = \langle Q_1, Q_3, Q_2 \rangle$

4. Stop, because $|R| = r$. 


Appendix

α-NDCG@k for a single query \( q \)

\[
\alpha - \text{NDCG}(q, k) = Z_{kq} \frac{G[m]}{\log_2(1+m)}
\]

- \( k \) - the number of results in the ranking that are considered
- \( Z_{kq} \) - a normalization factor such that the \( \alpha - \text{NDCG} \) of the perfect ranking for \( q \) at \( k \) is 1.

\[
G[m] = \sum_{i=1}^{S} J(d_m, n_i)(1 - \alpha)^r_{m-1} - \text{gain of a result}
\]

- \( J(d_m, n_i) \) - an assessor judged that document \( d_m \) contains nugget \( n_i \).
- \( \alpha \) - a factor that balances relevance and novelty
- \( r_{m-1} \) - the number of documents ranked up to position \( m - 1 \) that have been judged to contain nugget \( n_i \).
- \( S = \{n_1, ..., n_s\} \) - the set of information nuggets

**Jaccard Similarity between two queries \( Q_1, Q_2 \) and their interpretations \( (I_1, I_2) \)**

\[
\text{Sim}(Q_1, Q_2) = \frac{I_1 \cap I_2}{I_1 \cup I_2}
\]

**Query Score combining Relevance and Similarity**

\[
\text{Score}(Q) = \lambda \cdot P(Q|K) - (1 - \lambda) \cdot \sum_{q \in QI} \frac{\text{Sim}(Q, q)}{|QI|}
\]

- \( QI \) - a set of previously selected query interpretations in the ranked list
- \( \lambda \) - a factor to balance relevance and novelty in the diversification procedure

**Diversification Algorithm**

**Proc** Select Diverse Query Interpretations

1. copy the first element of \( L \) in \( R \);
2. while (less than \( r \) elements selected) {
   1. //select the best candidate for \( R[i] \)
   2. initialize best_score=0;
   3. while (more candidates for \( R[i] \) in \( L \)) {
      1. if (Score(\( L[j] \)) > best_score) set new best_score;
   }
   4. add the candidate from \( L \) with best_score to \( R \)
}

End Proc;