RAG Evaluation - MAP (Mean Average Precision)

📚 Given Data

Query 1

Relevant Documents:

- "Python is a programming language."
- "Python supports multiple libraries."

Retrieved Documents (Top-3):

Rank	Document	Relevant?
1	Python is a programming language.	✔ Yes
2	Python is a snake species.	✘ No
3	Python supports multiple libraries.	✔ Yes

Query 2

Relevant Documents:

- "Machine learning is a subset of AI."
- "Supervised learning uses labeled data."

Retrieved Documents (Top-3):

Rank	Document	Relevant?
1	Supervised learning uses labeled data.	✔ Yes
2	Unsupervised learning uses unlabeled data.	✘ No

Rank	Document	Relevant?
3	Machine learning is a subset of AI.	✔ Yes

📊 Step 1: Compute Average Precision (AP) for Each Query

Query 1

Calculate precision at each relevant document position:

Rank	Relevant?	Precision@Rank	Include in AP?
1	✔ Yes	1/1 = 1.000	☑ Yes
2	✖ No	-	✖ No (skip)
3	✔ Yes	2/3 ≈ 0.667	☑ Yes

Average Precision for Query 1:

$$\text{AP}_1 = \frac{1.000 + 0.667}{2} = \frac{1.667}{2} = 0.833$$

Query 2

Calculate precision at each relevant document position:

Rank	Relevant?	Precision@Rank	Include in AP?
1	✔ Yes	1/1 = 1.000	☑ Yes
2	✖ No	-	✖ No (skip)
3	✔ Yes	2/3 ≈ 0.667	☑ Yes

Average Precision for Query 2:

$$AP_2 = \frac{1.000 + 0.667}{2} = \frac{1.667}{2} = 0.833$$

🎯 Step 2: Compute MAP@3

Formula:

$$\text{MAP@3} = \frac{1}{N} \sum_{i=1}^{N} AP_i$$

where N = number of queries

Calculation:

$$\text{MAP@3} = \frac{AP_1 + AP_2}{2} = \frac{0.833 + 0.833}{2} = 0.833$$

✅ Final Results

Metric	Value
AP (Query 1)	0.833
AP (Query 2)	0.833
MAP@3	**0.833**

Interpretation

A MAP@3 score of **0.833** (83.3%) indicates strong retrieval performance. Both queries successfully retrieved relevant documents early in the rankings, with relevant items appearing at ranks 1 and 3 in both cases.