

ISR-3



Explain me the following concepts in easy understandable words

Following is the syllabus content of Unit 3 of Information Retrieval System subject:

Unit III Evaluation and Visualization of Information Retrieval System (06 hrs)

Performance evaluation: Precision and recall, MRR, F-Score, NDCG, user-oriented measures.

Visualization in Information System: Starting points, Query Specification, document context, User relevance judgment, Interface support for search process.



Here is a **simple, easy-to-understand explanation** of all concepts from **Unit 3 – Evaluation & Visualization of Information Retrieval Systems**.

★ **Unit III – Evaluation & Visualization of Information Retrieval Systems (Easy Explanation)**

1. Performance Evaluation of an IR System

Evaluation tells us **how good a search engine or retrieval system is.**

We use different metrics to measure accuracy and usefulness.

A) Precision

- Precision tells **how many retrieved results were actually relevant.**
- It checks **quality** of the results.



Example:

You searched for “Machine Learning notes”.
System shows 10 results, but only 6 are correct.
Precision = 6/10 = **0.6**

B) Recall

- Recall tells **how many relevant documents the system succeeded in finding.**
- It checks **coverage/completeness.**

Example:

Total correct documents in database = 20
System retrieved only 6
Recall = 6/20 = **0.3**

C) F-Score (F1-score)

- Combines **precision + recall** into one measure.
- Useful when you want a **balanced** performance.

Formula:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

D) MRR (Mean Reciprocal Rank)

- Measures **how early the first relevant result appears** in the ranking list.
- If a relevant document appears at Rank 1 → score = 1
Rank 2 → score = 1/2
Rank 5 → score = 1/5

MRR = average of all reciprocal ranks over many queries.

Why useful?



- Helps evaluate systems where **top-1** or **first correct answer** matters (e.g., Q&A systems).
-

E) NDCG (Normalized Discounted Cumulative Gain)

- Measures the **quality of ranked results**, considering:
 1. Highly relevant documents should appear early
 2. Lower relevant documents can appear later
- Gives higher score if **important documents appear at the top**.

Used by Google, YouTube, Netflix, etc.

F) User-oriented Measures

These evaluate the search process **from the user's perspective**, not just mathematically.

Examples:

- **Time taken** to find information
- **Number of clicks** needed
- **User satisfaction** rating
- **Ease of using the interface**
- **Navigation experience**

These measure **real-life usability** of the IR system.

2. Visualization in Information Retrieval Systems

Visualization helps users **see and understand search results better**.

It improves navigation, query refinement, and understanding of document relationships.

A) Starting Points

These are the **entry points** from where a user begins a search.

Examples:



- Search box
- Category list
- Previous search history
- Recommendations

Goal: make it easy for the user to start searching.

B) Query Specification

This is how the system helps a user **form and refine their search query**.

Includes:

- Suggesting keywords
- Auto-complete
- Spelling correction ("Did you mean...")
- Filters (date, file type, author)

Helps users express their need more accurately.

C) Document Context

Shows **where and how the search query appears in the document**.

Examples:

- Highlighting searched words
- Showing preview/snippet
- Showing surrounding text context
- Showing related documents

Helps users decide whether a document is relevant.

D) User Relevance Judgment

Users decide **which documents are relevant** to their query.

- Users may rate relevance (relevant / non-relevant)
- User behaviour such as clicks and time spent also indicate relevance
- Helps improve ranking algorithms



This feedback refines future search results.

E) Interface Support for the Search Process

Interface tools that make searching **easy, visual, and interactive**.

Examples:

- Filters and facets
- Sort options
- Graphical result views (charts, clusters)
- Document preview windows
- Pagination
- Recommendation panels

Goal: Help users find the best information with minimal effort.

🎯 Summary (Very Short & Simple)

Concept	Simple Meaning
Precision	How many retrieved were correct
Recall	How many correct were retrieved
F-score	Balance of precision and recall
MRR	How early the first correct result appears
NDCG	Quality of ranking based on relevance levels
User-oriented measures	Evaluating system based on user satisfaction & usability
Starting points	Where user begins searching
Query specification	Helping users frame a proper query
Document context	Showing matched keywords inside documents
User relevance judgement	User marking documents as relevant or not
Interface support	Tools that help users search better



If you want, I can also create **9-mark detailed answers, diagrams, or exam-ready notes** based on these.



Now give me proper ans of following que:

Define and explain following terms

- i)
 - ii)
 - iii)
- Precision
Recall
F-Score
[9]



Here is a **proper 9-mark answer** for the terms **Precision, Recall, and F-Score**, written in clear, exam-oriented language.

Q. Define and explain the following terms: (i) Precision, (ii) Recall, (iii) F-Score [9 Marks]

i) Precision

Definition:

Precision is a performance evaluation measure in Information Retrieval that indicates **the proportion of retrieved documents that are actually relevant** to the user's query. It focuses on the **accuracy or quality** of the results returned by the system.

Explanation:

A retrieval system may return many documents for a query, but not all of them are useful. Precision measures how many of the retrieved documents were correct.



Higher precision means the system retrieves **fewer irrelevant documents**, giving users more accurate results.

Formula:

$$Precision = \frac{\text{Number of Relevant Documents Retrieved}}{\text{Total Documents Retrieved}}$$

Example:

If the system retrieves 10 documents and only 7 are relevant,

Precision = $7/10 = 0.7$ (70% accurate results).

ii) Recall

Definition:

Recall is a measure that indicates **the proportion of all relevant documents in the database that the system successfully retrieves**. It focuses on the **completeness** of the search results.

Explanation:

Even if a system retrieves accurate results, it may still miss some relevant documents.

Recall tells us how effectively the system covers all relevant information.

Higher recall means the system retrieves **most of the relevant documents**, ensuring completeness.

Formula:

$$Recall = \frac{\text{Number of Relevant Documents Retrieved}}{\text{Total Relevant Documents in the Collection}}$$

Example:

If there are 20 relevant documents in the database and the system retrieves 7 of them,

Recall = $7/20 = 0.35$ (35% coverage).

iii) F-Score (F1-Score)

Definition:

The F-Score is a combined evaluation measure that balances **both precision and recall**. It represents the **harmonic mean** of precision and recall.



Explanation:

Precision and recall often conflict—improving one may reduce the other. The F-Score provides a single metric that considers both accuracy and completeness of retrieval. It is especially useful when a balanced evaluation of the IR system is required.

Formula:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Interpretation:

- High F-Score = system has **good accuracy and good coverage**
- Low F-Score = system fails on either precision, recall, or both

Example:

If Precision = 0.7 and Recall = 0.5,

$$F1 = 2 \times \frac{0.7 \times 0.5}{0.7 + 0.5} = 0.58$$

★ Final Answer Summary (Exam-Ready)

- **Precision** measures the accuracy of retrieved documents.
- **Recall** measures the completeness of retrieving all relevant documents.
- **F-Score** combines precision and recall into a single balanced measure using the harmonic mean.

If you want, I can also prepare answers for **MRR**, **NDCG**, or **Visualization concepts** in 8-9 mark format.



Calculate the precision and recall scores for the search, using the given data.

[6]

A database contains 160 relevant records on a particular topic and a



search was conducted on that topic and 120 records were retrieved. Out of 120 records retrieved, 90 were relevant



Here is the **step-by-step calculation** of **Precision** and **Recall** using the given data, written in clear exam format.

Q. Calculate the Precision and Recall scores [6 Marks]

Given:

- Total relevant records in database = **160**
 - Total records retrieved = **120**
 - Relevant records retrieved = **90**
-

1. Precision

Formula:

$$Precision = \frac{\text{Relevant Retrieved}}{\text{Total Retrieved}}$$

Substitute values:

$$Precision = \frac{90}{120}$$

$$Precision = 0.75$$

Precision = 0.75 or 75%

2. Recall

Formula:

$$Recall = \frac{\text{Relevant Retrieved}}{\text{Total Relevant in Database}}$$

Substitute values:

$$Recall = \frac{90}{160}$$

$$Recall = 0.5625$$

Recall = 0.5625 or 56.25%

✓ Final Answers:

- **Precision = 0.75 (75%)**
- **Recall = 0.5625 (56.25%)**

If you want, I can also calculate **F-Score** or prepare this in a **stepwise solved-paper format**.



Explain in detail the term NDCG. Explain with a suitable example.[6]



Below is a **clear, exam-oriented 6-mark answer** for **NDCG**, with definition, explanation, working, and a simple example.

★ NDCG (Normalized Discounted Cumulative Gain) (Definition + Explanation + Example) [6 Marks]

Definition:

NDCG (Normalized Discounted Cumulative Gain) is a **ranking quality measure** used in Information Retrieval to evaluate how well a system ranks documents based on their **relevance levels**.



It gives **higher scores** when highly relevant documents appear **at the top positions**, and lower scores when relevant items appear later in the list.

Why NDCG is Needed?

Simple precision or recall only check *whether* relevant documents appear, but not **where** they appear.

Example:

- A highly relevant document at **rank 1** is far more useful than at **rank 7**.
NDCG captures this importance by **discounting relevance with rank**.
-

★ Components of NDCG

1) DCG (Discounted Cumulative Gain)

DCG measures the total usefulness (gain) of retrieved documents, giving **more weight to top-ranked results**.

Formula:

$$DCG = rel_1 + \frac{rel_2}{\log_2(2)} + \frac{rel_3}{\log_2(3)} + \dots$$

Where **rel_i** = relevance score of document at position *i*
(Relevance can be 0, 1, 2, 3 based on importance.)

Higher ranks = less discount → more score.

2) IDCG (Ideal DCG)

IDCG is the **best possible ranking** where documents are perfectly sorted from most relevant to least relevant.

It is used to normalize the score.

3) NDCG Calculation

$$NDCG = \frac{DCG}{IDCG}$$



NDCG values lie between **0 and 1**, where

- **1 = perfect ranking**
 - **Closer to 0 = poor ranking**
-

🌟 Example (Easy to Understand)

Assume a search engine returns results with the following **relevance scores**:

Rank	Relevance
1	3
2	2
3	0
4	1

Let's compute NDCG@4

Step 1: Compute DCG

$$DCG = 3 + \frac{2}{\log_2(2)} + \frac{0}{\log_2(3)} + \frac{1}{\log_2(4)}$$

Now calculate:

- $\log_2(2) = 1$
- $\log_2(4) = 2$

$$DCG = 3 + \frac{2}{1} + \frac{0}{1.58} + \frac{1}{2}$$

$$DCG = 3 + 2 + 0 + 0.5 = 5.5$$

Step 2: Compute IDCG (Ideal Ranking)

Sort relevance scores from highest → lowest:

3, 2, 1, 0

$$IDCG = 3 + \frac{2}{\log_2(2)} + \frac{1}{\log_2(3)} + \frac{0}{\log_2(4)}$$



Calculate:

$$IDCG = 3 + 2 + \frac{1}{1.58} + 0$$

$$IDCG = 3 + 2 + 0.63 = 5.63$$

Step 3: Compute NDCG

$$NDCG = \frac{5.5}{5.63} \approx 0.98$$

✓ Final Interpretation

- **NDCG = 0.98**, meaning the system's ranking is **almost perfect**, very close to the ideal ranking.
 - Most relevant documents appear at the top positions—good ranking quality.
-

Summary (Short Notes for Exam)

- NDCG evaluates ranking quality considering **position + relevance level**.
 - Uses **DCG** to measure real ranking and **IDCG** to measure ideal ranking.
 - Produces a normalized score between **0 and 1**.
 - Higher NDCG = better ranking performance.
-

If you want, I can also prepare **MRR**, **User-oriented measures**, or the entire **Unit 3 notes** in exam answer format.



Define and explain Interface support for search process related to visualization in information system





Below is a **clear, exam-oriented answer** for “**Interface Support for Search Process**” related to visualization in Information Systems.

★ Interface Support for Search Process

(*Definition + Explanation*) — *Visualization in Information Systems*

Definition:

Interface support for the search process refers to the **visual and interactive tools provided by the search interface** that help users perform a search effectively. It includes all the **features, controls, visual cues, and navigation aids** that assist users in expressing their query, exploring results, refining searches, and understanding document relevance.

★ Explanation in Detail:

In an Information Retrieval System, the user interacts with the system through an interface.

A well-designed interface greatly **improves the user experience** by making it easier to:

- Enter and refine queries
- View results clearly
- Compare documents
- Navigate through large result lists
- Understand which documents are most relevant

Visualization techniques and interactive components play a major role in this support.

★ Key Components of Interface Support

1. Search Box and Query Assistance



The interface helps users form queries by providing tools like:

- Auto-complete suggestions
- Query expansion
- Spell-correction (“Did you mean...”)
- Synonym suggestions

This reduces user effort and improves query quality.

2. Filters and Faceted Search

Users can refine results using filters such as:

- Date
- Author
- File type
- Category

Facets help users narrow down results visually and interactively.

3. Result Presentation and Visualization

The interface visually displays search results using:

- Snippets or previews
- Highlighting of matched keywords
- Document summaries
- Icons, thumbnails, or category tags

This helps users quickly **judge relevance**.

4. Sorting and Ranking Controls

Users can change the order of results by:

- Relevance
- Date
- Popularity



- File size

Allows flexible and personalized exploration.

5. Interactive Navigation Tools

Includes:

- Pagination
- Infinite scrolling
- Back/forward navigation
- Breadcrumbs

These help users move easily through large sets of results.

6. Relevance Feedback Options

The interface may allow:

- Marking documents as “relevant or not relevant”
- Clicking, rating, or saving documents
- Personalized recommendations

This feedback improves the ranking algorithm.

7. Document Comparison and Context Views

Interfaces may offer:

- Side-by-side comparison
- Context view (where query terms appear in the document)
- Highlighting occurrences

This helps users make better relevance decisions.

⭐ Importance of Interface Support

Interface support **directly affects**:

- User satisfaction



- Speed of finding information
- Accuracy of selecting relevant documents
- Efficiency of the overall search process

Good visualization makes complex data understandable and improves retrieval quality.

✓ Exam-Ready Summary (Short)

Interface support for the search process includes all visual and interactive features of a search interface that help users conduct a search smoothly.

It provides tools for query formulation, result visualization, filtering, navigation, relevance judgment, and document context viewing.

Effective interface support improves user experience, reduces effort, and enhances the overall performance of the information retrieval system.

If you want, I can also create a **6-mark answer** or **pointwise short answer** for your exam.



What are User oriented measures used in performance evaluation of IR systems.



Below is a **clear, exam-ready answer** for **User-Oriented Measures** used in performance evaluation of Information Retrieval (IR) systems.

⭐ User-Oriented Measures in Performance Evaluation of IR Systems

User-oriented measures evaluate an Information Retrieval system **from the user's point of view**, rather than only looking at mathematical accuracy.



These measures focus on **user satisfaction, search experience, usability, and effort required** to find the correct information.

Unlike precision and recall, user-oriented measures consider how real users interact with the system and how effectively the system supports their needs.

★ Key User-Oriented Measures

1. Response Time / Search Time

How quickly the system returns results.

Shorter response time means a **better user experience** and higher satisfaction.

2. User Satisfaction

Users may give ratings such as:

- How useful were the results?
- Was the information easy to find?
- How well did the system meet their expectations?

Surveys or questionnaires are used to measure this.

3. Number of Clicks to Find Relevant Information

Measures **how much effort** the user must put in.

Fewer clicks indicate a more efficient search process.

4. Ease of Use / Usability

Evaluates how simple the interface is:

- Easy navigation
- Clear result display
- Good organization of filters and options

A system that is easy to use increases retrieval efficiency.



5. Relevance Judgment by Users

Users judge documents as:

- Relevant
- Partially relevant
- Not relevant

These judgments help understand how well the system matches user expectations.

6. Learning Curve

Measures **how quickly a new user can understand and use the system**.

A low learning curve means the system is intuitive.

7. Session-level Measures

Looks at the entire search session instead of a single query:

- How many queries user issued
- How many refinements made
- Whether user found final answer successfully

This reflects real-world search behavior.

8. User Engagement

Includes:

- Time spent on relevant documents
- Document browsing patterns
- Bookmarking or saving results

Higher engagement often means better relevance.

★ Summary (Short Answer for Exams)



User-oriented measures evaluate IR systems based on **user experience and usability**.

They include measures such as **response time, user satisfaction, number of clicks, ease of use, relevance judgments, learning curve, engagement, and session-level search effectiveness.**

These measures help understand how well the system supports users in finding the right information with minimum effort.

If you want, I can also create **3-mark, 6-mark, or 9-mark** versions depending on your exam pattern.



Define Query. What are various techniques used to specify query in information visualization?



Here is a **clear, exam-oriented answer** for the question:

Define Query. What are the various techniques used to specify query in information visualization?

★ Definition of Query

A **query** is a *user's request for information* from an Information Retrieval (IR) system. It represents what the user wants to find, usually expressed through keywords, phrases, filters, or visual interactions.

A query guides the system to search and retrieve the most relevant documents or data from the database.



⭐ Techniques Used to Specify Queries in Information Visualization

Information visualization provides several interactive and graphical methods to help users **formulate, refine, and express queries** easily.

These techniques make query specification more intuitive than typing text alone.

1. Direct Manipulation

Users interact directly with visual elements like:

- Clicking
- Dragging
- Selecting
- Highlighting

Example: Selecting a range on a timeline or zooming into a graph to filter data.

2. Form-Based Query Specification

Users fill forms containing:

- Text fields
- Dropdowns
- Checkboxes
- Filters

Common in advanced search interfaces (e.g., filter by date, author, category).

3. Keyword-Based Query

The traditional method where users type keywords in a search box.

Often enhanced by:

- Auto-complete
 - Spell correction
 - Query suggestions
-



4. Query-by-Example (QBE)

Users choose an example item, and the system retrieves similar items.

Example: Selecting a document or image and asking the system for "similar results."

5. Visual Query Specification

Users express queries through **graphs, maps, charts, or diagrams**.

Examples:

- Drawing a shape to filter data
 - Using sliders to select numerical ranges
 - Highlighting clusters to view items inside them
-

6. Dynamic Queries

Queries update results **in real time** as the user adjusts controls such as:

- Sliders
- Checkboxes
- Buttons

Example: Moving a slider to filter data by date, instantly updating results.

7. Faceted Search

Users refine queries using multiple facets such as:

- Category
- Price
- Color
- Rating
- Location

Common in e-commerce and digital libraries.

★ Summary (Exam Ready)



A **query** is a user's request to retrieve specific information from an IR system. In information visualization, queries can be specified using techniques such as **direct manipulation, form-based input, keyword search, query-by-example, visual queries, dynamic queries, and faceted search.**

These techniques make the searching process more intuitive, interactive, and visually clear for the user.

If you want, I can convert this into a **3-mark, 4-mark, or 6-mark** shorter version.



Define and explain following terms
i)
ii)
iii)
MRR
NDCG
E-measure [9]



Below is a **proper 9-mark, exam-oriented answer** for **MRR, NDCG, and E-measure**, explained clearly with definitions and examples.

★ **Define and Explain the Following Terms:

- (i) MRR
 - (ii) NDCG
 - (iii) E-measure [9 Marks]**
-

i) MRR (Mean Reciprocal Rank)

Definition:



MRR is a performance evaluation measure used in Information Retrieval to determine **how early the first relevant result appears** in the ranking returned by the system.

Explanation:

- If a relevant document appears at **Rank 1**, the reciprocal rank = 1
- If it appears at **Rank 3**, reciprocal rank = 1/3
- If no relevant document is found, the reciprocal rank = 0

MRR is the **average** of reciprocal ranks computed over multiple queries.

Formula:

$$MRR = \frac{1}{N} \sum_{i=1}^N \frac{1}{rank_i}$$

Where $rank_i$ is the rank of the first relevant document for query i .

Example:

For three queries, relevant documents appear at ranks: 1, 4, 2

$$MRR = \frac{1}{3} \left(1 + \frac{1}{4} + \frac{1}{2} \right) = \frac{1.75}{3} = 0.58$$

Higher MRR → system returns correct answers earlier.

ii) NDCG (Normalized Discounted Cumulative Gain)

Definition:

NDCG is a ranking quality metric that evaluates **how well a system orders documents based on their relevance levels**, giving higher weight to relevant documents appearing at higher ranks.

Explanation:

- It considers **graded relevance** (e.g., 0, 1, 2, 3)
- Uses logarithmic discounting so that documents at top ranks contribute more



- Normalizes the score by comparing it with the **ideal** ranking

NDCG values range between **0 and 1**, where 1 means a **perfect ranking**.

Steps:

1. **DCG (Discounted Cumulative Gain):**

$$DCG = rel_1 + \frac{rel_2}{\log_2(2)} + \frac{rel_3}{\log_2(3)} + \dots$$

2. **IDCG (Ideal DCG):**

Best possible ordering of relevance values.

3. **NDCG:**

$$NDCG = \frac{DCG}{IDCG}$$

Example:

Relevance at ranks = [3, 2, 0, 1]

$$DCG = 3 + \frac{2}{1} + \frac{0}{1.58} + \frac{1}{2} = 5.5$$

Ideal ranking = [3, 2, 1, 0]

$$IDCG = 3 + 2 + 0.63 = 5.63$$

$$NDCG = \frac{5.5}{5.63} = 0.98$$

High NDCG → top results are well ranked.

iii) E-measure

Definition:

E-measure is a combined performance metric that evaluates **both precision and recall simultaneously**.



It is similar to the F-measure but allows adjusting the balance between precision and recall.

Explanation:

- E-measure represents the **error** in retrieval performance
 - Lower E-measure indicates a **better** retrieval system
 - It includes a weighting parameter (β) to emphasize precision or recall as needed
 - If $\beta > 1 \rightarrow$ recall is given more importance
 - If $\beta < 1 \rightarrow$ precision is given more importance
-

Formula:

$$E = 1 - F_\beta$$

Where

$$F_\beta = \frac{(1 + \beta^2) \cdot P \cdot R}{(\beta^2 \cdot P) + R}$$

β = weight parameter, P = precision, R = recall

Thus,

$$E = 1 - \frac{(1 + \beta^2)PR}{\beta^2 P + R}$$

Interpretation:

- E = 0 \rightarrow best performance
 - E near 1 \rightarrow poor performance
-

Example:

If precision = 0.75, recall = 0.60, $\beta = 1$:

$$F_1 = \frac{2 \cdot 0.75 \cdot 0.60}{0.75 + 0.60} = 0.666$$

$$E = 1 - 0.666 = 0.334$$



Lower E-measure means good retrieval quality.

★ Final Summary (Exam-Ready)

Term	Meaning	Purpose
MRR	Mean Reciprocal Rank	Measures how early the first relevant document appears
NDCG	Normalized Discounted Cumulative Gain	Measures ranking quality considering relevance levels and position
E-measure	Error-based combined metric	Balances precision & recall; lower value = better performance

If you want, I can prepare **short 3-mark versions** or **full Unit-3 long answers** for practice.



Why are the performance evaluation measures needed in IR system?
Explain trade-off between Recall and Precision.
[6]



Below is a **clear, exam-oriented 6-mark answer** for:

“Why are performance evaluation measures needed in IR systems? Explain the trade-off between Recall and Precision.”

★ Why are Performance Evaluation Measures Needed in IR Systems?

Performance evaluation measures are essential in Information Retrieval (IR) systems because they help determine **how effectively the system retrieves relevant**



information.

They provide a scientific and objective way to analyze the strengths and weaknesses of a retrieval system.

Reasons why evaluation measures are needed:

1. To measure accuracy and relevance

Evaluation metrics like Precision, Recall, F-score, etc., tell how many retrieved documents are actually useful and how many relevant documents are being missed.

2. To compare different IR systems or algorithms

Metrics allow researchers to compare different search algorithms and choose the best-performing one.

3. To improve system performance

Evaluation reveals weaknesses such as high irrelevant results or missing important documents.

Developers can use this information to improve ranking algorithms.

4. To ensure user satisfaction

Users expect a search system to return accurate, relevant, and quick results.

Performance measures help judge whether the system meets user expectations.

5. To test real-world usefulness

Evaluation measures show how well the system behaves for real queries, ensuring reliability in practical usage.

⭐ Trade-off Between Recall and Precision

Precision and Recall often behave **oppositely**, meaning improving one may reduce the other.

This is known as the **Precision-Recall Trade-off**.



Recall

- Measures how many relevant documents are retrieved
 - High recall = more coverage
 - To increase recall → system retrieves **more documents**, including some irrelevant ones
-

Precision

- Measures how many retrieved documents are actually relevant
 - High precision = high accuracy
 - To increase precision → system retrieves **fewer but highly relevant documents**
-

★ Why is there a Trade-off?

1. Increasing Recall decreases Precision

If the system retrieves *more documents* to ensure no relevant ones are missed, it will also pick up **irrelevant documents**, lowering precision.

2. Increasing Precision decreases Recall

If the system retrieves *only highly accurate documents*, it may ignore some relevant documents, lowering recall.

★ Example to Show the Trade-off

Suppose a system retrieves only 5 highly relevant documents:

- Precision = High
- Recall = Low (because many relevant documents remain unretrieved)

If it retrieves 100 documents:

- Recall = High
 - Precision = Low (because irrelevant ones are included)
-



⭐ Conclusion

Precision and Recall must be balanced depending on the application.

- Medical diagnosis → High Recall (don't miss important cases)
 - Legal search or recommendation → High Precision (only high-quality results)
-

If you want, I can also prepare a **diagram** showing the trade-off or convert this answer into a **short 4-mark version**.

