

## IR Question Bank

Short answer questions.

1. Define the term IR. State the objectives of IR Systems.
2. What are advantages and disadvantages of Information Retrieval.
3. Explain four key elements associated with IR model.
4. What are problems with Boolean retrieval models?
5. Define following terms:
  - i) TF – IDF
  - ii) Relevance feedback
  - iii) Cosine similarity
6. What are the steps involved in text pre-processing?
7. How Does Query Expansion Work?
8. What is Zipf's law in NLP?
9. What are applications of Text clustering ?
10. What are levels of text clustering?
11. What are Recommendation Systems?
12. K-Dimensional Tree (KDTree) is an improvement over KNN. Justify your answer.
13. Differentiate between : Stemming and Lemmatization
14. What are the uses of Naive Bayes classifiers with respect to IR systems?
15. Explain K means clustering algorithm.
16. Explain the factors that influence the PageRank algorithm.
17. Define : a) Query expansion (QE)
  - b) Web Crawler
18. What do you mean by Metasearch engines? Give eg.
19. Define: a) Levenshtein distance
  - b) Jaccard Similarity
  - c) Hamming distance
- 20: Define : a) overstemming  
b)understemming  
c)Inverted Index

## 5 Marks questions

1. What is IR model? Explain Boolean Model.
2. Explain: Different text similarity algorithms
3. Explain four key elements associated with IR model.
4. What do you mean by stopwords? How to remove stopwords?
5. What is Query Expansion?What are applications of Query Expansion?
6. Write short note on: Text Clustering
7. Explain: Hierarchical clustering
8. Explain the working of Search Engine with diagram.
9. What is Expectation-Maximization Algorithm?
10. Explain: K-Dimensional Tree (KDTree)
11. Explain the working of :Naive Bayes Classifiers
12. What Is Collaborative Filtering?
13. Explain: Content-based recommendation system
14. What is difference between Explicit Feedback and Implicit Feedback with reference to user query improvement.
15. What is crawling with respect to search engine?
16. What is Search engine indexing?
17. What is a robots.txt file?
18. Explain : Porter stemmer algorithm
19. Explain : Snowball stemmer algorithm
20. Explain: Online measures of evaluation of IR system.

## 10 Marks questions

1. What is IR Model? Explain four key elements associated with IR model. Explain Vector Space Model.
2. Write Short notes on:
  - i) K means clustering algorithm
  - ii) Parts of IR system
3. Explain : K-Nearest Neighbor(KNN) Algorithm
4. What is Stemming and Lemmatization? Support your answer with proper examples.  
Explain types of Stemmer in NLTK
5. Write Short notes on:
  - i) Agglomerative clustering
  - ii) Zipf's law in NLP
6. Explain the working of Search Engine with proper diagram.Explain the importance of PageRank algorithm.

7. What is Relevance Feedback with respect to user query improvement?
8. What is web crawling? Explain the process of crawling with the architecture diagram.
9. Explain the Components of Information Retrieval model with diagram.
10. What are types of IR models? Explain any 2.
11. Short notes:
  - a) Tokenization
  - b) Text Normalization
12. What are offline metrics for evaluation of IR system? Explain with eg.