ISR Oral Question Bank

Assignment: 1

1.	What is Conflation Algorithm? What Is Need Of Conflation?
2.	What is Luhn's Idea?
3.	What are Different Methods Of Stemming?
4.	What is Diff Algorithm For Suffix Stripping
5.	What is Major in Porter's Algorithm?
6.	What is Document Representative?
7.	What are The Contents/Formats Of Doc Representative?
	Ans:(Keyword(Conflated(Stemmed)),Frequency, Occurrences, Position)
8.	How much % The Doc Reduces After Conflation?
9.	What is Use of Finding Similarity
10.	What are Different Technique Need for Finding Similarity
11.	What is Rule for Removing 'ing' In Porters Algorithm?
12.	Different Stages of Conflation.
13.	Calculate precision and recall for following system Different examples like database with 34 record ,50 retrieved 30 are relevant 20 are irrelevant
14.	How to measure performance of IR system.
15.	Alternative measure for evaluation of IR System
16.	Upper Bound and lower-Bound in Luhn's idea

Assignment 2:

17.	What is Clustering?
18.	What is Need of Centroid?
19.	What are the Differences Between Clustering and Classification?
20.	What are The Different Methods Of Clustering? Which Method you have used?
21.	What are Different Similarity Functions/Measures?
22.	Write an Example of Clustering.
23.	What is Cluster Hypothesis?(Pg 30)
24.	What is Dendogram?
25.	What are Different Search Strategies?(Unit 5)
26.	What is the Difference Between Single Pass Algorithm And Single Link Algorithm?(Pg 37)
27.	What is Measure of Association?(Pg 24, 25)
28.	What are the Different Methods of Classification?(Pg 28)
29.	What are the Use of Clustering in IR?
30.	What is Minimum Spanning Tree? Where It is used?
31.	

Assignment 3:

32.	What is indexing?
33.	What is the use of indexing?

34.	What is the need of indexing?
35.	What are the data structure used in indexing'?
36.	What are different types of files?(pg 49)
37.	How you will search in indexing?(pg 66 67)
38.	What is dictionary?
39.	What is posting list?
40.	Implement the assignment till the end!
41.	What is index term weighting?
42.	What is probalistic index?
43.	Why the index is called as inverted index?
44.	Difference in all 3 IR models.
45.	Explain Probabilistic Model

.

Assignment 4:

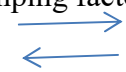
46.	What is crawling?
47.	When does Google use web crawler?
48.	Difference between browsing and crawling?
49.	Use of scheduler and crawler
50.	What is browsing?
51.	Why we use scheduler in crawler?
52.	What is multithreaded downloader in crawler?
53.	What is role of crawler?
54.	What is URL?
55.	What is meta crawler?
56.	What is meta searcher?
57.	What are methods of crawling?
58.	How BFS and DFS is used in crawler?

Assignment 5:

59.	What is diff between 2D image and 3D image?
60.	How the image is read?
61.	What are the content of image?
62.	What re-pixel of color image represent?
63.	What is feature extraction of image?
64.	Which are property uses for image extraction?
65.	What are image contents?
66.	How the identify content in image?
67.	How color image is represented?
68.	What is size of color image matrix?
69.	(every pixel represent with 3 value and is 0 to 255)
70.	How pixel represent in gray scale image?
71.	What are other feature can we extracted from 2D image?
72.	What is content based image retrieval?

73.	What is multimedia based image retrieval?
74.	How image is represented in memory?

Assignment 6:

75.	What is web search engine?
76.	What is lexicon in Google search engine?
77.	What is difference between search engine and IR?
78.	Which algorithm used by Google page ranking.
79.	What is difference between Typical IR and web search as an IR?
80.	Lucene Architecture.
81.	Lucene block diagram and functioning of various blocks.
82.	What is web search engine?
83.	What is lexicon in Google search engine?
84.	What is page ranking? Calculate the page rank of following web pages assume dumping factor 0.7 A  B  C
85.	Discuss challenges involved in web searching?

Assignment 7:

86.	What is Semantic web and recommended system?
87.	Component of recommended system.
88.	Designing recommender system what all components can be used?
89.	Difference between IR and Recommender system?
90.	What is function, features and properties of recommender system?
91.	Draw typical block diagram of search engine.
92.	Difference between normal computer search and IR System?
General questions:	
93.	How we evaluate IR system?(ch 7 pg 112)
94.	What is relevance?
95.	What is relevance feedback?
96.	What are different model for information retrieval? (pg 21)
97.	What is information system?
98.	What is data retrieval and information retrieval?
99.	Explain Exhaustively??
100.	What is distributed IR? (ch 9 neto)?
101.	What is suffix tree and suffix array?
102.	What is signature file?
103.	What are different algo for sequential search?
104.	What is File in Google?

105.	Benefits and needs of automatic classification?
106.	What is Semantic web?
107.	Probabilistic model of information retrieval.
108.	What is ontology?
109.	What is page ranking? How page ranking work? Explain Google page ranking?
110.	IF 100 documents/books are given to you how you will make a IR system
	out of it so that wherever you search, it will take least amount of time.
111.	If different clothes of various types are given to you for your shop, how
	you will arrange it so that you will be able to serve the customer in
	minimum amount of time? (Hint: Classification)
112.	What is the difference between windows search and Google's search?
113.	If you have been a task to make a IR system for your college student's
	Information, which information (available, new) will be used ?
114.	What do you think data retrieval system should replace by IR system?
115.	What is the difference between structured and unstructured data?
116.	Where are they used? Give Example.
117.	List diff tools for data analysis and text analysis?
	Answer : R, WEKA, Map Reduce, Spark, Lexica, etc
118.	What are the examples of IR, Data Retrieval?
119.	How will you measure relevance of the terms in the documents? (Hint :
	TF, IDF)
120.	Function of collaborative system? Types? Examples?
121.	How the KNN Algorithm is used and for which purpose?
122.	What are diff web query languages?
123.	Explain K-means algorithm for clustering.
124.	Draw arch. for content based recommender system.
125.	Example for Content based and item based