**Exam.Code:0920**
**Sub. Code: 6819**

**1019**

**B.E. (Computer Science and Engineering)**

**Eighth Semester**

**Elective – V**

**CS-803C: Information Retrieval and Management**

**Time allowed: 3 Hours** **Max. Marks: 50**

*NOTE: Attempt five questions in all, including Question No. I which is compulsory and selecting two questions from each Unit.*

x-x-x

I. Attempt the following:-

a) Give some mechanism for web size measurement.

b) Explain the decision tree tor classification

c) What is parametric search?

d) Explain Tf-ldf scoring.

e) Explain any spelling correction method. (5x2)

**UNIT – I**

II. a) What is inverted index? Draw the inverted index that would be built for the following document collection:
Dl: new home sales top forecasts, D2: home sales rise in July, D3: increase in home sales in July, D4: July new home sales rise.

b) Differentiate stemming and lemmatization by explaining the functioning of Porter stemmer and any lemmatizer. Also explain the use of positional indices. (5,5)

III. a) What do you understand by wild-card queries? Explain in detail the two ways of handling the wild-card queries.

b) Give a dynamic programming based algorithm for calculating edit distance. (5,5)

IV. a) Give and explain the Blocked Sort-Based Indexing by giving example.

b) Differentiate dynamic indexing and n-gram indexing by quoting suitable examples. (5,5)

**UNIT – II**

V. a) Give and explain the algorithm for computing the weighted zone score from two postings lists.

b) Explain in detail by giving examples and uses of Vector Space Model based scoring. (5,5)

P.T.O.

VI. a) What do you understand by reduced dimensionality reductions? Give their use in computing scores in a search system.

b) Give the Naive Bayes classification approach by mentioning all the mathematical details. (5,5)

VII. a) Explain the Support Vector Machine classifiers by giving all the inner details.

b) Explain the functioning of any focused crawler. (5,5)

*x-x-x*