

# SPPU-BE-COMP-CONTENT – KSKA Git

Total No. of Questions : 4]

SEAT No. :

PE-193

[Total No. of Pages : 2

[6580]-553

**B.E. (Computer Engineering) (Insem)**

**INFORMATION RETRIEVAL**

**(2019 Pattern) (Semester - VII) (Elective- IV) (410245 (A))**

*Time : 1 Hour]*

*[Max. Marks : 30*

*Instructions to the candidates:*

- 1) *Solve questions Q.1 or Q.2, Q.3 or Q.4.*
- 2) *Neat diagrams must be drawn wherever necessary.*
- 3) *Figures to the right indicate full marks.*
- 4) *Assume suitable data if necessary.*

**Q1) a)** Define Information Retrieval. Demonstrate probabilistic model with example. **[7]**

b) Explain TF-IDF Model in detail with example. **[8]**

Consider 3 documents in corpus

Doc1: "Taj Mahal is a beautiful monument"

Doc2: "Red Fort is also a monument"

Doc3: "I like to visit Agra"

Draw the term document incidence matrix for this document collection.

OR

**Q2) a)** Explain stemming and tokenization with example. **[7]**

b) Demonstrate vector space model with example. **[8]**

PT. <sup>P.T.O.</sup>

# SPPU-BE-COMP-CONTENT – KSKA Git

Q3) a) What are the different types of index construction and explain disk based index construction with suitable examples. [7]

b) Write a short note on: [8]

i) Term-at-a-time query processing

ii) Hash-based dictionary

OR

Q4) a) Explain Lightweight Structure and Generalized Concordance Lists [7]

b) Explain different classes of operators used in Query optimization [8]

