(3 Hours)   [Total Marks : 80]

Please check whether you have got the right question paper.

N.B.:   1)   Question No.1 is compulsory
        2)   Attempt any three questions from remaining.
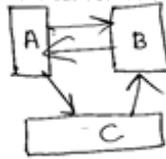   3)   Assume suitable data if necessary.

1. Answer the following
    a) Define Edit distance. Explain with example.   **(05)**
    b) What is the role and effect of page rank?   **(05)**
    c) List the different issues in stream processing.   **(05)**
    d) Distinguish between Name node and Data node.   **(05)**

2. a) Compare and contrast   **(10)**
        i)   Multistage and Multihash
        ii) SQL and NoSQL
    b) Explain FM algorithm with example?   **(10)**

3. a) Compute simplified page rank using damping factor d = 0.9 for web.   **(10)**

    b) Explain selection and projection relational algebraic operation using   **(10)**
        MapReduce.

4. a) Explain the different Hadoop components?   **(10)**
    b) Explain SIMRANK Algorithm with suitable example?   **(10)**

5. a) Define Hub and Authority. Compute Hub and Authority scores for web.   **(10)**

    b) Explain different types of recommendation system with real time examples? **(10)**

6. Answer the following   **(20)**
    a) NoSQL data stores
    b) SON algorithm with Map Reduce
    c) MapReduce for Matrix-Multiplication
    d) Clustering using Representatives Algorithm

_____