

## JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

## TEST -2 EXAMINATION- 2016

## B.Tech VIII Semester

COURSE CODE: 15B1WCI833

MAX. MARKS: 25

COURSE NAME: Big Data Analytics

COURSE CREDITS: 03

MAX. TIME: 1Hr 30 Min

---

*Note: All questions are compulsory. Carrying of mobile phone during examinations will be treated as case of unfair means.*

---

1. What do you understand by stream data? Give real life examples. What are the problems to process the stream data? How do you differentiate the standing queries from ad hoc one? Differentiate with suitable examples. [5]
2. Compare the sampling of a fixed proportion with sampling of a fixed-size sample for sampling from a data stream. Prove reservoir sampling with principal of mathematical induction. [5]
3. Prove that the space complexity of DGIM method which is used to count the number of 1s in the recent stream tuples is  $O(\log^2 N)$ . Exercise DGIM method to count the number of 1s in the stream 101010000001010100000100001111100001111100000010101010101 [5]
4. Explain different types of parameters used to quantify importance of a web page across the internet with suitable mathematical formulas. [5]
5. Explain the working of search engine with vector space model. Take suitable example to illustrate your answer. Take suitable example to illustrate the disadvantage of the ordinary working of a search engine. Suggest method to overcome the said problem. [5]