MCA
(SEM V) THEORY EXAMINATION 2017-18
BIG DATA

Time: 3 Hours   *Total Marks: 100*

Notes: *Attempt all Sections. Assume any missing data.*

Section A

1. **Attempt all parts:**   (10X2=20)

   a) What is Big Data? Explain with an example.
   b) Define the various file formats supported by HIVE.
   c) Compare & Contrast NOSQL relational Database.
   d) List four disadvantages of aggregate Oriented Database.
   e) What is Map reducing?
   f) State the purpose of Hadoop pipes.
   g) How the pig programs can be packaged?
   h) List down the entities of YARN.
   i) Specify the role of job tracker & task tracker in HDFS.
   j) What is wrapper class and list out the wrapper classes in java?

Section B

2. **Attempt any five parts.**   (10X5=20)

   a) Explain with neat sketch about the processing of a job in hadoop.
   b) Explain the role of driver code, mapper code and reducer code with in a map reduce program model by a suitable example.
   c) Explain the Master slave and peer-peer replication in detail.
   d) Explain about the various data types supported by pig in its data model with an example.
   e) Explain about various data types supported by HIVEQL with an example.
   f) List and discuss various dimensions of Big data.
   g) What do you mean by serialization and how should be the RPC serialization format?
   h) Explain the architecture of HIVE with neat sketch.

Section C

Attempt any two parts.   (15X2=30)

3. Distinguish between the old and new versions of Hadoop API for Map Reduce frame work.
4. How do you analyze Big Data? Explain in detail with industry examples.
5. Explain these terms with an example
   a) Cassandra Data Model
   b) Crowd Sourcing Analytics