**SESSION 1: HOW TO SOLVE BIG DATA**

**ASSIGNMENT 2**

**1. What constitutes Big Data?**

a.) volume

b.) velocity

c.) variety

**d.) All the above**

**2. Solution Identified to handle Big Data is?**

**a.) Hadoop**

b.) storm

c.) spark

d.) All of these

**3. Doug cutting added DFS & MapReduce support to:**

**a) Nutch**

b) Spark

c) GFS

d) Facebook

**4. Which file system is followed by Hadoop?**

a.) DFS

b.) GFS

**c.) HDFS**

d.) NTFS

**5. How much data can be considered as a BigData with regards to volume?**

a.) Mega Bytes

b.) Terabytes

c.) Petabytes

**d.) b & c**

**6. Can Hadoop process data in real time?**

**a.) Yes**

b.) No

c.) None of these

**7. Hadoop is developed using?**

a.) C

b.) C++

**c.) Java**

d.) Ruby

**8. Hadoop reduces cost of operation by?**

**a.) Using commodity Hardware**

b.) Using less number of nodes

c.) Limited cluster size

d.) None of these

**9. Search engine adopted in Hadoop is?**

a.) Google

b.) Yahoo

**c.) Solr**

d.) AOL

**10. Stream processing can be achieved in Hadoop using?**

a.) Storm

b.) Spark

**c.) Both of these**

d.) None of these

**11. What is the highest data unit we have at present?**

a.) Petabyte

b.) Exabyte

c.) Terabyte

**d.) Brontobyte**

**12. Hadoop is well suited for clusters where:**

**a.) Number of nodes keep changing**

b.) Number of nodes are constant

c.) Most number of free nodes

d.) None of these

**13. What is batch processing?**

**a.) processing of previously collected jobs in a single batch.**

b.) processing of previously collected jobs in different batches.

c.) processing of jobs at that moment they are collected

d.) None of these

**14. What is stream processing?**

**a.) Processing data as streams**

b.) Processing data in batches

c.) Both of these

d.) None of these

**15. Hadoop uses?**

a.) Batch processing

b.) Near Real time processing

c.) Real time processing

**d.) All the above**