

**Tutorial 2**

1. Explain the importance of YARN and HDFS in Hadoop architecture.
2. There are 5 main pillars in Hadoop. Give one example tool for each pillar and explain the functionality of the tool.
3. By taking a real-life big data example, explain how the data pipeline for the particular big data solution is designed.
4. By taking a real-life example, explain how machine translation works, and compare it with human translation, in terms of efficiency and accuracy.
5. Discuss the advantages and disadvantages to use autocoding method in big data resources.
6. Given a big data resource that contains text about cognitive science, explain how term extraction is carried out to create index automatically for the resource.