Assignment-4

Sai Hemanth Nirujogi
G01065588

A high potential for success is locked away in the big data of every organization. By hitting on the right spot strategically, the organization can mount a huge advantage, from being productive to being profitable. A deeper understanding of the big data that is being generated by the organization every day can provide better visibility and insights into the future of the organization, which in turn helps the organization in making better decisions, engage in better planning and better agility to counter growing market and competition. [1]

To realize the true potential of the big data, and to address the key challenges to an effective analytics solution like an estimated price, speed, precision and expertise, there are two ways:

- Build/DIY
- Buy

Characteristics of Big data efforts that might require the need for a "Build / Do It Yourself" approach

- To deliver a complete view of the organization's business model, which completely justifies all the organization's data, both structured and unstructured from a wide variety of sources.
- To check if the data is suitable for the purpose every single time, so the data can be utilized for the benefit of the organization.
    - Accuracy
    - Finished
    - Consistency
- To deliver a complete view of the organization's information to all the stakeholders, and other individuals regardless of their position or level of expertise in the company.
- To maximize the effectiveness of the big data and utilize its potential to the fullest in providing insights and trends considering all the departments like the business, IT etc., in the organization. [2]
- To operationalize the data trends, and creating insights for the business users in the organization to view the effective patterns and identify the predictions.

Examples of potential commercial ("Buy") products that failed to meet extreme data analytics requirements

- **SAS**

  SAS is a programming language and a statistical tool which is used to handle huge amounts of data generated every single day. Its main job is to structure and analyze the data. This can provide data from unstructured format to readable format. SAS is preferred by every large organization right now for big data analytics and as a language for analysis, regardless of its price tag. Even though it is the global leader in big data analytics it lacks graphic representation of data, which is one of the most important things to analyze and to find trends in the data. Text mining can be a real pain with SAS compared to a more traditional choice like R. Text mining helps an organization in making decisions from extracting information from the big data. SAS is a closed environment and not open source, which can be a problem for small organizations. [3]

- **IBM Watson Analytics**

  IBM Watson is a supercomputer which integrates the power of Artificial Intelligence and Analytics for a greater performance machine. The key components of the Watson include Apache's Hadoop, SUSE Linux Server, 2880 processors, 15 terabytes of RAM and IBM's DeepQA software. The features and power of Watson are incredible and endless with text mining and performing really complex analytics on massive volumes of unstructured data. The commercial implementation of Watson has first happened in 2013, for Memorial Sloan Kettering Cancer Center started using Watson to suggest Lung cancer treatments to the patients with low costs. Watson Analytics is also an effective implementation of its technology to explore, visualize and to present the data by using all the power and features of Watson. [4] With all this power and capabilities Watson cannot process the structured data directly and provide insights. With also limited amounts of resources while the data is growing is hard to maintain with Watson. IBM has made the AI as the main component of the Watson rather than using cognitive computing power. The features and possibilities once offered by Watson are now offered by Azure, AWS and Data Ninja for a cheaper price, the only thing that is keeping Watson running in business is the branding once it has. [5]

References

1. “Building End-to-End Big Data Solutions (Part 1).” *MapR Media Kit | MapR*, accessed on 29th September 2018 mapr.com/blog/building-end-end-big-data-solutions-part-1/.
2. “Should You Build or Buy Your Big Data Analytics Platform?” *Datameer,* accessed on 29th September 2018 www.datameer.com/blog/build-buy-big-data-analytics-platform/.
3. “A Complete SAS Tutorial – Learn Advanced SAS Programming 2018.” *DataFlair*, accessed on 29th September 2018 data-flair.training/blogs/sas-tutorial/.
4. “What Is IBM Watson Supercomputer? - Definition from WhatIs.com.” *SearchEnterpriseAI,* accessed on 30th September 2018 searchenterpriseai.techtarget.com/definition/IBM-Watson-supercomputer.
5. Brown, Jennings. “Why Everyone Is Hating on IBM Watson-Including the People Who Helped Make It.” *Gizmodo*, Gizmodo.com, 10 Aug. 2017, accessed on 30th September 2018 gizmodo.com/why-everyone-is-hating-on-watson-including-the-people-w-1797510888.