2017

Big Data and Analytics

CONCEPTS, METHODS, TECHNIQUES, AND APPLICATIONS

KIMBERLY ROETEN

CIS 445-01 | Data Mining

Big data and the concepts of analytics are topics that have been around for years, and have also proven to be essential and beneficial to the success of most companies today. For years, they were basic numbers on a spreadsheet that businesses could get significant value from, now they can be used to uncover certain insights that can potentially improve a company's efficiency and their decision-making process. Big data analytics can now be used as a competitive advantage because they give organizations the capability of working at a much faster pace.

Big data and analytics have become very important in the business world for purposes of identifying new favorable opportunities that in turn could lead to an increase of customer satisfaction, higher profits, and more efficient business measures. Some of the data that has been captured by most organizations includes the ideas of cost reduction, faster and better decision making, and new goods and services. When it comes to storing large amounts of data, analyzing information immediately and determining customer needs by creating new products prove to be some of the most helpful tools to improving an organization's efficiency.

Big Data is such a big topic that it impacts just about every industry and organization that ultimately depends on understanding the sources of data. Not only does big data itself come from the increases and improvements of data storage, but also from growing streams that every individual in a business contributes. Because the subject itself is so overwhelming and can be unmanageable if not properly understood, some organizations have trouble with making the best value of every part of it. To make it not seem so massive, the best solution to make it more understandable is to illustrate the concepts, tools, and applications that drives the topic itself. Breaking it down into smaller pieces such as where it comes from, what makes it big, how it can potentially be leveraged, and future directions is a step. Managing, storing, and computing big

data and analytics is not a one-step process, so analyzing valuable opportunities and moving away from the somewhat usual way of doing things can give organizations the kind of personalization it needs, especially when it comes to investment.

Most of the time, when people stop to look at the term “Big Data”, for sure a marketing term comes to mind, but what most people don’t understand is big data and analytics are important for advancing trends in technology such as new approaches that could lead to better decisions. When I think of the term, I think of the movie “Get Smart” in a way where the chief explains to Max that he spends half his time trying to convince the Vice President that intelligence comes from human beings and not satellites. True, the intelligence of certain technologies can be applied in many business fields, but most of the time is artificial. We as humans make the data more accessible and most importantly more understandable to computers, and if we don’t understand it to begin with, leaving computers in charge of all the crucial decisions will only hurt business efficiency.

Big data and analytics provide a means of analyzing immense data sets and ultimately drawing conclusions from them to help organizations make more informed business decisions. Unlike regular data analytics, big data analytics is more advanced so to speak, where the statistical data behind it is powered more by higher performance systems. It is so advanced that it requires more complex applications that can involve predictive models, algorithms, and other mainstream analyses. In 2001, analyst Doug Laney increased the concept of big data to not only refer to increasing data volumes, but to also increase the variety of data being generated by organizations and the velocity where data is created and updated on a daily basis. So volume, variety, and velocity then became the three factors that were popularized in 2005.

What most people don't know is that big data types typically don't fit well in traditional data warehouses because they are typically made for structured relational data sets, and big data aims more towards unstructured and semi-structured data types. Because of this, big data types are not able to handle processing commands that may need to be updated on a daily basis, so as a result, many companies that analyze big data have turned to NoSQL databases so the collection and processing are more manageable so to speak. Many of its companion tools include Hadoop, YARN, Spark, and HBase, and since their discovery they have been used as staging areas for big data and analytics before they are put into the final data warehouse. In Hadoop, data is analyzed either in a cluster or in a processing engine, and these processes serve as a storage unit for incoming streams of raw data. And data that is stored must be in an organized fashion, and configured to receive good performance and feedback. When data is officially ready, it is analyzed with the software that is used for an organization's analytics processes. This is where the idea of data mining comes in. Data mining tools come in handy in the analytics processes because they examine the data sets for certain patterns and relationships which then lead to the process of building models used to predict customer needs and behavior, and potentially future developments that can improve this. Other than data mining, text mining and statistical analysis also play a vital role in big data and analytics, as these types of software can mainstream data visualization tools crucial for analytics applications.

So far I have pointed out the uses of big data and analytics, however there are certain challenges that come with them due to the fact that many of the applications are pulled from both internal and external sources, especially where data is being produced by a third-party service provider. The amount of data that is typically involved can potentially cause data management issues in areas including data quality and consistency, and if these areas are not kept up to date

and managed carefully, this can result in the use of different data platforms which can lead to potential disasters that can cause organizations to lack certain analytics skills. And when the lack of experience increases, this forces organizations to hire more experienced data scientists and engineers to fix the problems, which will only increase the overall cost, and one of the points of data management is to reduce cost.

Even though the organization of big data and analytics is a decision made by the management team most of the time, every single developer part of the organization must buy into the certain values it brings to the system and the company as a whole. Ending up with worthless and unutilized data will bring major risks to the table. Establishing stable supporters and mentors is a first step in enhancing analytics. Even though management may play a huge role in how quickly organizations will adapt to these platforms, it is important to keep in mind that data itself is never central, meaning decisions that are made will eventually reach all layers of a company, and if employees are not willing to participate in providing feedback that helps business make solid decisions, the success of analytics is questionable. As Daniel Newman said in his forbes article, “it’s not what you know, it’s what you do with what you know.” Companies sooner or later must realize that the proper and effective use of analytics is what drives the majority of business operations.

Many organizations have faced overwhelmingly large amounts of data, it’s a wonder the term itself is called Big Data. Due to the complexity and organizations rapidly changing over the years from the pressure of its competitors, doors full of possibilities have opened up to improve customer needs. For those that use big data and analytics, the rates of productivity and profitability no doubt become higher based on the data advantage that is pulled from wanting to develop more promising destinations.

Citations:

https://www.sas.com/en_us/insights/analytics/big-data-analytics.html#

https://www.forbes.com/sites/mckinsey/2013/07/22/big-data-analytics-and-the-future-of-marketing-sales/#14362fd85587

http://www.predictiveanalyticsworld.com/newyork/2017/big_picture.php

http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html

http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics

https://www.ibm.com/analytics/us/en/technology/hadoop/big-data-analytics/

https://www.forbes.com/sites/danielnewman/2017/01/31/realizing-the-potential-of-big-data-and-analytics/#5ede19d5709a