ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ

ΤΜΗΜΑ ΔΙΟΙΚΗΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΕΙΡΗΝΗΣ ΣΟΦΙΑΝΟΥ

BIG DATA AND PREDICTIVE ANALYTICS IN SUPPLY CHAIN, FOCUS ON PREDICTIVE MAINTENANCE AND E-COMMERCE

Επιβλέπων : Ι.ΜΟΥΡΤΟΣ ,Αναπληρωτής Καθηγητής

Υποβληθείσα ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος (MSc) στη Διοικητική Επιστήμη και Τεχνολογία

Αθήνα, Φεβρουάριος 2019

Βεβαίωση εκπόνησης Διπλωματικής εργασίας

«Δηλώνω υπεύθυνα ότι η συγκεκριμένη μεταπτυχιακή εργασία για τη λήψη του μεταπτυχιακού τίτλου σπουδών του ΠΜΣ στη Διοικητική Επιστήμη και Τεχνολογία του Τμήματος Διοικητικής Επιστήμης και Τεχνολογίας του Οικονομικού Πανεπιστημίου Αθηνών έχει συγγραφεί από εμένα προσωπικά και δεν έχει υποβληθεί ούτε έχει εγκριθεί στο πλαίσιο κάποιου άλλου μεταπτυχιακού ή προπτυχιακού τίτλου σπουδών στην Ελλάδα ή το εξωτερικό. Η εργασία αυτή έχοντας εκπονηθεί από εμένα, αντιπροσωπεύει τις προσωπικές μου απόψεις επί του θέματος. Οι πηγές στις οποίες ανέτρεξα για την εκπόνηση της συγκεκριμένης διπλωματικής αναφέρονται στο σύνολό τους, δίνοντας πλήρεις αναφορές στους συγγραφείς, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο».

(Υπογραφή)

....

ΣΟΦΙΑΝΟΥ ΕΙΡΗΝΗ

Φοιτητής MSc στη Διοικητική Επιστήμη και Τεχνολογία

Abstract

Big data, which means a large amount of data, has numerous sources. These meters persistently stream data about electricity, water, or gas utilization that can be shared with customers and combined with valuing plans to motivate customers to move some of their energy utilization. Collecting big data is not enough though, as today, the overwhelming reason for this ineffective management is the absence of truthful data to evaluate the genuine need for repair or maintenance of plant machinery, equipment, and systems. Maintenance scheduling has been, and in numerous instances still is, predicated on statistical trend data or on the genuine failure of plant equipment. Based on the above definitions, this paper analyzes the correlation between supply chains, big data, data analytics, predictive maintenance and e-Commerce.

Table of Contents

1. Introduction

In recent years big data and predictive analytics (BDPA) has been at the center of dialogs among academics and industry. With the revolutions in technology, data are generated a lot faster and in higher volume than at any other time (Duan and Xiong, 2015). Big data is characterized by 3 basic parts: volume, velocity and variety (Zhou et al. 2014). On the basis of fundamentals of data science, we can argue that data analytics benefits from large volume of data. The statistical reliability tends to increase with the increase in volume of data (i.e. populace size increases) (in the same place). Furthermore, predictive analytics' methods with more number of variables have better explanatory power than one or few components. Velocity refers to the rate at which data are generated. Today due to online sales, advanced cells, social networks and sensory devices, the flow of data has increased fundamentally. Variety refers to the different types of data that are available, that is, unstructured, semi-structured, and structured data. Hence, we can argue that big data can possibly contribute to the predictive analytics in two different ways: high reliability and high explanatory power. Duan and Xiong (2015) have noted that BDPA can possibly revolutionize the existing supply chains. BDPA can be extensively used for enhancing supply chain performance by enhancing visibility which is identified as one of the most essential organizational capability to improve organizational performance (Barratt and Oke, 2007) and improve resilience and robustness (Brandon-Jones et al. 2014). Columbus (2015) characterizes BDPA as a capability that generates cost reserve funds for SCM processes and contributes to the competitiveness of a firm. Other scholars underline the importance of BDPA for enhancing organizational performance (OP) (Schoenherr and Speier-Pero, 2015), leveraging decision-production (Bose, 2006), and changing the supply chain (Jeyraj et al., 2006; Waller and Fawcett, 2013). McGuire et al. (2012) further argue that innovative firms seek to beat competition by finding better approaches to leverage BDPA for next-generation items and services, increasing data transparency and decision-production effectiveness by means of data digitization and accessibility, and precisely segmenting their customer base as per the 'who', 'what', 'when', and 'where' for different items and services. Therefore, BDPA helps with achieving higher levels of performance (Waller and Fawcett, 2013).

In the previous few years, an explosion of interest in big data has occurred from both academia and the e-commerce industry. This explosion is driven by the way that e-commerce firms that inject big data analytics (BDA) into their value chain experience 5– 6% higher profitability than their competitors (McAfee and Brynjolfsson, 2012). A recent report by BSA Software Alliance in the United States (USA) indicates that BDA contributes to 10% or more of the development for 56% of firms (Columbus, 2014). Therefore, 91% of Fortune 1000 companies are investing in BDA projects, an 85% increase from the previous year (Kiron et al., 2014a). While the use of emerging internet-based technologies provides e-commerce firms with transformative benefits (e.g., real-time customer service, dynamic evaluating, personalized offers or improved interaction) (Riggins, 1999), BDA can further set these effects by enabling informed decisions based on basic bits of knowledge (Jao, 2013). Specifically, in the e-commerce context, "big data enables merchants to track each user's behavior and connect the spots to determine the best approaches to convert one-time customers into repeat buyers" (Jao, 2013, p.1). Big data analytics (BDA) enables e-commerce firms to use data more efficiently, drive a higher conversion rate, improve decision making and empower customers (Miller, 2013). From the perspective of exchange cost theory in e-commerce (Devaraj et al., 2002; Williamson, 1981), BDA can benefit online firms by enhancing market exchange cost efficiency (e.g., buyer-seller interaction online), managerial exchange cost efficiency (e.g., process efficiencyrecommendation calculations by Amazon) and time cost efficiency (e.g., searching, bartering and after sale checking). Drawing on the resource-based view (RBV) (Barney, 1991), we argued that BDA is a distinctive competence of the elite business process to 3 bolster business needs, for example, identifying faithful and profitable customers, determining the ideal price, detecting quality problems, or deciding the lowest possible level of inventory (Davenport and Harris, 2007a). Notwithstanding the RBV, this research views BDA from the relational metaphysics of sociomaterialism perspective, which advances the argument that different organizational capabilities (e.g., management, technology and talent) are constitutively entangled (Orlikowski, 2007) and commonly supportive (Barton and Court, 2012) in achieving firm performance. At last, service marketing offers the perspective of enhancing service development models, which has been reflected by firms, for example, Rolls Royce (Barrett et al., 2015), Amazon, Google and Netflix (Davenport and Harris, 2007a). In that capacity, the extant literature identifies BDA as the stage for "development of employment, increased efficiency, and increased consumer excess" (Loebbecke and Picot, 2015, p.152), the "next big thing in advancement" (Gobble, 2013, p.64); "the fourth worldview of science" (Strawn, (2012); "the next frontier for advancement, competition, and profitability" (p. 1) and the next "management revolution" (p. 3) (McAfee and Brynjolfsson, 2012); or that BDA is "getting a revolution science and technology" (Ann Keller et al., 2012); etc. Due to the high effect in e-commerce, eminently in generating business value, BDA has recently become the focal point of academic and industry investigation (Fosso Wamba et al., 2015c). In spite of the fact that an increasing measure of published materials has focused on practitioners in this space, the literature remains largely anecdotal and fragmented. There is a scarcity of research that provides a general scientific classification from which to explore the dimensions and utilizations of big data in e-commerce.

2. Personal involvement to the subject

Logistics and any research around its sector, has for more than 4 years been an interesting subject to be involved in. After my three years engagement with a bit more classic, strict form of logistics , that analyzing data and planning according to analytics was considered as a waste of time and capital,I finally found my way to a more “open minded” perspective of logistics planning.

With my new team, we aim to pivot the old fashioned ways of the already existing logistics corporations in Greece. To achieve our goal, we daily analyze and report our existing clients, keeping big ammounts of data so as to make better transportation, warehousing and palleting planning. Moreover, economic offers to potential customers are extracts of a deep research on customers’ historical and behavioural data , so as to achieve the best of quality and price correlation.

My basic and daily responsibilities are centred on checking IOD (Information on Demand) from our daily shipments and make correction when needed, analyzing the delivery times and rapidly interfere when significant discrepancies occur. Furthermore, I double check the RFID status in all delivery cycle, so as to be valid from storage to delivery and attest that all data flows are correctly saved and categorizes in order to be further analyzed.

This new work role, actually gave me more impetus to make a thesis for big data and all the new trends of this particular sector. Even if, analyzing big data may prerequisites more technical background, understanding the meaning and importance of big data and modern data analyzing methods, is actually the key to pivoting coorporations in Greece.

3. Data Analytics in Supply Chain

3.1. Theoretical Framework

During the last decades, many new technical management tools have been released and took the management processes to a whole new level. New ideas based on selecting and processing the already available data in corporations, led us to a new business world, where data analysis is considered as a science with millions of supporters. New terminologies, such as Big Data, Data Mining and Predictive Analytics, have been deeply discussed, analyzed and presented in many Business Universities throughout the world.

Even though, 90% of the available data has been created and collected the last four years and Big Data came out on 2005 when the O'Reilly Media was launched, the meaning of data has been around much longer.

More and more companies tended to shift to e-sectors in the late 90s,while understanding that the data that accrued from their commerce transactions could be used in a more productive way, than just being archived in the premises of the corporations.

Moingeon and Edmonson, in early 1996, referred a simple example of the volume of data produced in few minutes in e-commerce businesses just by basic customers' transactions.

Even if the data was getting more and more important in businesses procedures, the term Big Data, as a large set of data almost impossible to be managed and processed using traditional business tools, was coined to Roger Mougalas in 2005. Web 2.0 led to major increase of the data created on daily basis, which triggered the activity of innovative startups to dig into massive amount of data. Governments and big corporations, from this moment on, will spend major amounts of capital and effort, to cope with the changes and stay in the game. Big Data is considered to be "as far as the internet was in 1993" , a fact that clearly shows that Big Data revolution is still ahead of us, in all of social and business sectors.

From an evolutionary perspective, big data isn't new. A noteworthy reason for creating data warehouses during the 1990s was to store large measures of data. In those days, a terabyte was considered big data. Teradata, a leading data warehousing vendor, used to recognize customers when their data warehouses reached a terabyte. Today, Teradata has more than 35 customers, for example, Wal-Mart and Verizon, with data warehouses over a petabyte in size. eBay captures a terabyte of data per minute and keeps up over 40 petabytes, a large portion of any organization on the planet. So what is big data? One perspective is that big data is more and different sorts of data than is easily handled by customary relational database management systems (RDBMSs). Some people consider 10 terabytes to be big data, however any numerical definition is likely to change over time as associations collect, store, and analyze more data. Another useful perspective is to characterize big data as having high volume, high velocity, and high variety—the three Vs [Russom, 2011]:

High volume—the sum or amount of data

High velocity—the rate at which data is created

High variety—the different types of data

So, "big data" means there is more of it, it comes more rapidly, and comes in more structures. Both of these perspectives are reflected in the accompanying definition [Mills, Lucas, Irakliotis, Rappa, Carlson, and Perlowitz, 2012; Sicular, 2013]: Big data is a term that is used to describe data that is high volume, high velocity, and/or high variety requires new technologies and techniques to capture, store, and analyze it and is used to enhance decision making, provide understanding and discovery, and bolster and optimize processes.

It is helpful to recognize that the term analytics isn't used consistently it is used in no less than three different yet related ways [Watson, 2013a]. A beginning stage for understanding analytics is to explore its underlying foundations. Decision emotionally supportive networks (DSS) during the 1970s were the primary systems to help decision making [Power, 2007]. DSS came to be used as a description for an application and an academic discipline. Over time, extra decision bolster applications, for example, executive data systems, online investigative processing (OLAP), and dashboards/scorecards became well known.

Then during the 1990s, Howard Dresner, an investigator at Gartner, popularized the term business intelligence. A normal definition is that "BI is a general category of uses, technologies, and processes for gathering, putting away, accessing, and investigating data to help business users make better decisions" [Watson, 2009a, p. 491]. With this definition, BI can be viewed as an umbrella term for all applications that help decision making, and this is the way it is interpreted in industry and, increasingly, in academia. BI evolved from DSS, and one could argue that analytics evolved from BI (at any rate in terms of terminology). Along these lines, analytics is an umbrella term for data investigation applications. BI can likewise be viewed as "getting data in" (to a data shop or warehouse) and "getting data out" (examining the data that is stored). A second interpretation of analytics is that it is the "getting data out" some portion of BI. The third interpretation is that analytics is the use of "rocket science" calculations (e.g., machine learning, neural networks) to analyze data. These different takes on analytics don't ordinarily cause much perplexity, because the context as a rule makes the meaning clear (A. Katal, M. Wazid, R. H. Goudar 2013).

3.2. Big Data

Understand that what is believed to be big data today won't seem so big in the future. Numerous data sources are currently untapped—or if nothing else underutilized. For example, every customer e-mail, customer-service visit, and social media comment might be captured, stored, and analyzed to better understand customers' sentiments. Web perusing data may capture every mouse movement with the end goal to better understand customers' shopping behaviors.

Radio frequency identification (RFID) labels might be placed on every single piece of merchandise with the end goal to assess the condition and area of every item.

Big data has numerous sources. For example, every mouse tap on a web site can be captured in Web log files and analyzed with the end goal to better understand shoppers' purchasing behaviors and to influence their shopping by powerfully recommending items. Social media sources, for example, Facebook and Twitter generate tremendous measures of comments and tweets. This data can be captured and analyzed to understand, for example, what people consider new item presentations. Machines, for example, shrewd meters, generate data. These meters persistently stream data about electricity, water, or gas utilization that can be shared with customers and combined with valuing plans to motivate customers to move some of their energy utilization, for example, for washing clothes, to non-peak hours. There is a tremendous measure of geospatial (e.g., GPS) data, for example, that created by cell phones, that can be used by applications like Four Square to help you know the areas of friends and to receive offers from nearby stores and restaurants. Image, voice, and sound data can be analyzed for applications, for example, facial recognition systems in security systems (A. Katal, M. Wazid, R. H. Goudar 2013).

3.3. Applications in Supply Chains

A supply chain has been defined as "a bidirectional flow of data, items and money between the underlying suppliers and last customers through different associations" SCM includes arranging, implementing and controlling this flow. In the current computerized economy, SCs have been viewed as key levers for competitive advantage. That is presumably why some scholars argue that competition inside the market space has evolved from "firm versus firm" towards "supply chain versus supply chain" . In this context, the reception and use of innovative IT has been considered as a basic resource for SC streamlining. For example, SCM assumes a vital role in limiting an organization's overall danger of misrepresentation, bribery, and defilement. Earlier studies identified numerous benefits related to IT-enabled SC streamlining, including end-to-end data sharing among SC stakeholders, intra-and inter-organizational business process change (e.g., cancellation, redesign, robotization), improved decision-production inside the SC, improved operational efficiency, and increased revenue. BDA is expected to take SC change to a level of change never before achieved. For example, BDA represents a basic source of meaningful data that may help SC stakeholders to increase improved bits of knowledge they can use for competitive advantage. Likewise, BDA could help SC stakeholders to reduce their exposure to different dangers including the danger of extortion and other malfeasance. In the context of SC execution, BDA could lead to increased efficiency and gainfulness in the SC by boosting speed and visibility, enhancing SC stakeholders' relationships, and enhancing SC readiness.

BDA could result in faster time to market and the potential for superior revenue recognition. However, some investigators argue that the deluge of data threatens to "break the existing data supply chain» (Andrew K.S. Jardine, Daming Lin, and Dragan Banjevic 2006).

Settling on well-informed decisions in context involves a wide range of supply chain operations—from demand sensing and forecasting of inventory arranging and logistics wanting to execution and warehouse management, just to name a few. Presently more than any time in recent memory data sources are plenteous and in different structures and builds extending from GPS data to enable unique steering and scheduling of deliveries, purpose of sales (POS) data, operational data of warehouses, generation line data, inventory data and a wide range of types of structured and unstructured data from numerous parties over the entire logistics network. The fundamental point behind the preference of big data over little data is to uncover hidden patterns and experiences utilizing large sources of structured and unstructured data that would be obscured whenever limited selective and little data sources were used. Truth be told, whenever used correctly, big data holds the key to enhancing supply chain development by ensuring data integrity, and increased visibility and control through the supply anchor to increase nimbleness and responsiveness. Putting excessively emphasis on dissecting data sources that don't benefit supply chains can waste resources. Then again, expending excessively few, making it impossible to explore the sources of data available to basic supply chain operations may result in lost chance and cause unexpected interruptions to supply chain operations, and an unsuitable customer experience.

Moreover, big data can be a powerful apparatus for driving supply chains forward. Big data arrangements have helped large retail supply chains screen customer behavior and make more accurate predictions of customer preferences. Whether it's Walmart's capacity to anticipate the surge in demand for Pop Tarts amid a hurricane that enabled them to stock up in time , Amazon's capability to commence fulfillment and transportation processes before customers even place merchandise in advanced shopping baskets or eBay's capacity to identify the type of web design that will yield the most sales . Despite the potential provided by big data, many supply chains report that they're unable to harness the power of available data to generate useful bits of knowledge for their businesses.

The underlying reasons reveal a loss of motion by investigation due to supply chains either deficient with regards to the capabilities to analyze large totals of data or the correct segment of data being erroneous and causing the supply affix to bring about huge costs without achieving the desired outcomes. To maintain a strategic distance from loss of motion supply chains need to correctly identify potential problems and the correct set of data. Next, considering the sum and type of data needed to solve the problem, the correct type of analyses, abilities and apparatuses required must be selected to enable the problem to be tackled in the most cost effective manner. In the accompanying sections, we will shed light on how progressive supply chains are investing in data and analytics (D&A) capabilities (Andrew K.S. Jardine, Daming Lin, and Dragan Banjevic 2006).

Despite the largest development of data analytics being experienced in downstream customer bits of knowledge, analytics can have applications over the end-to-end supply chain Supply chains that are embracing big data capability development, first need to become aware of the benefits that big data arrangements can deliver to their operations. Decisions need to be made about the cost effectiveness of organizing certain parts of their operations. All-encompassing big data arrangements applied to the whole supply chain can involve high costs, settling on supply chain decision makers more selective in altering answers for specific operations. Below are mentioned some usages of big data analytics applications in the supply chains.

Increased visibility of inventory levels, demand, and assembling limit hence more accurate creation and dispersion scheduling

Real-time leading of big data examination inside the warehouse ERP system and identifying inventory levels, delivery miss-matches, and approaching deliveries

More accurate estimation of demand by accessing data of sales, market trends, competitors' data, and relevant nearby and worldwide economic components

Checking of delivery routes, movement data, weather in real-time and rerouting if necessary for limit and asset sharing

Full transparency at the SKU level and completely automated replenishment systems combined with demand forecast data that eliminate under/overstocking and optimize inventory running

Real-time enhancement of complex webs of conveyance centers, plants, and warehouses based on the material flow data

4. Predictive Maintenance

4.1. Definitions

Maintenance costs are a noteworthy piece of the aggregate operating costs of all assembling or creation plants. Depending on the specific business, maintenance costs can represent between 15 and 60 percent of the cost of merchandise produced. For example, in foodrelated industries, average maintenance costs represent around 15 percent of the cost of merchandise produced, whereas maintenance costs for iron and steel, mash and paper, and other heavy industries represent up to 60 percent of the aggregate creation costs. These percentages might be misleading. In most American plants, reported maintenance costs include numerous nonmaintenance-related expenditures. For example, numerous plants include adjustments to existing capital systems that are driven by market-related components, for example, new items. These expenses are not really maintenance and ought to be allocated to nonmaintenance cost centers however, true maintenance costs are significant and do represent a fleeting improvement that can directly affect plant gainfulness. Recent surveys of maintenance management effectiveness indicate that 33%—33 cents out of every dollar—of all maintenance costs is wasted as the result of unnecessary or improperly carried out maintenance. When you consider that U.S. industry spends more than $200 billion each year on maintenance of plant equipment and facilities, the effect on profitability and benefit that is represented by the maintenance operation becomes clear.

The overwhelming reason for this ineffective management is the absence of truthful data to evaluate the genuine need for repair or maintenance of plant machinery, equipment, and systems. Maintenance scheduling has been, and in numerous instances still is, predicated on statistical trend data or on the genuine failure of plant equipment. Up to this point, middle-and corporate-level management have ignored the effect of the maintenance operation on item quality, generation costs, and more imperative, on primary concern benefit. The general assessment has been "Maintenance is a necessary evil" or "There is no hope to improve maintenance costs." Perhaps these statements were true 10 or 20 years prior, however the development of microprocessor-or computerbased instrumentation that can be used to screen the operating state of plant equipment, machinery, and systems has provided the means to manage the maintenance operation. This instrumentation has provided the means to reduce or eliminate unnecessary repairs, prevent cataclysmic machine failures, and reduce the negative effect of the maintenance operation on the productivity of assembling and creation plants (D. Pfisterer, et al., 2009).

4.2. Types of Maintenance

Run-to-Failure Management

The rationale of run-to-failure management is simple and clear: When a machine breaks down, settle it. The "On the off chance that it ain't broke, don't settle it" method of keeping up plant machinery has been a noteworthy piece of plant maintenance operations since the primary assembling plant was manufactured, and on the surface it sounds reasonable. A plant utilizing run-to-failure management does not spend any money on maintenance until the point when a machine or system neglects to operate. Run-to-failure is a reactive management technique that sits tight for machine or equipment failure before any maintenance move is made however, it is really a "nomaintenance" approach of management. It is additionally the most expensive method of maintenance management. Few plants use a true run-to-failure management rationality. In all instances, plants perform fundamental preventive undertakings (i.e., grease, machine adjustments, and other adjustments), even in a run-to-failure environment. In this type of management, however, machines and other plant equipment are not rebuilt, nor are any significant repairs made until the point that the equipment neglects to operate. The real expenses associated with this type of maintenance management are high spare parts. An Introduction to Predictive Maintenance inventory cost, high overtime work costs, high machine downtime, and low generation accessibility (D. Pfisterer, et al., 2009) .

Preventive Maintenance

There are numerous definitions of preventive maintenance, yet all preventive maintenance management programs are time-driven. In other words, maintenance assignments are based on elapsed time or long periods of operation. The mean-time-to-failure (MTTF) or bath curve indicates that a new machine has a high likelihood of failure because of establishment problems amid the initial few weeks of operation. After this underlying period, the likelihood of failure is relatively low for an extended period. After this typical machine life period, the likelihood of failure increases forcefully with elapsed time. In preventive maintenance management, machine repairs or rebuilds are scheduled based on the MTTF measurement. The genuine implementation of preventive maintenance varies greatly. Some projects are extremely limited and comprise of just grease and minor adjustments. Comprehensive preventive maintenance programs schedule repairs, grease, adjustments, and machine rebuilds for all basic plant machinery. The shared factor for these preventive maintenance programs is the scheduling guideline—time.

Predictive Maintenance

Like preventive maintenance, predictive maintenance has numerous definitions. To some workers, predictive maintenance is monitoring the vibration of turning machinery trying to detect incipient problems and to prevent calamitous failure. To others, it is monitoring the infrared image of electrical switchgear, motors, and other electrical equipment to detect developing problems. The normal premise of predictive maintenance is that regular monitoring of the real mechanical condition, operating efficiency, and other indicators of the operating state of machine-prepares and process systems will provide the data required to ensure the most extreme interval between repairs and minimize the number and cost of unscheduled outages created by machine-train failures. An Introduction to Predictive Maintenance is Typical bath curve. Predictive maintenance is substantially more, however. It is the means of enhancing profitability, item quality, and overall effectiveness of assembling and generation plants. Predictive maintenance isn't vibration monitoring or thermal imaging or greasing up oil investigation or any of the other nondestructive testing techniques that are being marketed as predictive maintenance tools. Predictive maintenance is a theory or attitude that, basically stated, uses the real operating state of plant equipment and systems to optimize total plant operation. A comprehensive predictive maintenance management program uses the most costeffective tools (e.g., vibration monitoring, thermography, tribology) to acquire the real operating state of basic plant systems and based on this real data schedules all maintenance activities on an as-needed basis. Counting predictive maintenance in a comprehensive maintenance management program optimizes the accessibility of process machinery and greatly reduces the cost of maintenance.

It additionally improves the item quality, efficiency, and gainfulness of assembling and generation plants. Predictive maintenance is a condition-driven preventive maintenance program. Instead of relying on modern or in-plant average-life measurements (i.e., mean-time-to-failure) to schedule maintenance activities, predictive maintenance uses direct monitoring of the mechanical condition, system efficiency, and other indicators to determine the real mean-time-to-failure or loss of efficiency for each machine-train and system in the plant. Best case scenario, customary time-driven methods provide a guideline to "ordinary" machine-train life ranges. An official conclusion in preventive or run-to-failure programs on repair or rebuild schedules must be made on the basis of instinct and the personal experience of the maintenance manager.

Predictive maintenance utilizing vibration signature investigation is predicated on two fundamental actualities: (1) all regular failure modes have unmistakable vibration frequency components that can be isolated and identified, and (2) the amplitude of each particular vibration component will remain steady unless the operating elements of the machinetrain change. These certainties, their effect on machinery, and methods that will identify and evaluate the underlying driver of failure modes are developed in more detail in later chapters.

Total Productive Maintenance

Touted as the Japanese way to deal with effective maintenance management, the TPM concept was developed by Deming in the late 1950s. His concepts, as adapted by the Japanese, stress absolute adherence to the fundamentals, for example, oil, visual inspections, and universal use of best practices in all aspects of maintenance. TPM isn't a maintenance management program. A large portion of the activities associated with the Japanese management approach are directed at the generation work and assume that maintenance will provide the essential errands required to keep up basic creation assets. The majority of the quantifiable benefits of TPM are couched in terms of limit, item quality, and total creation cost. Unfortunately, domestic advocates of TPM have tried to implement its concepts as maintenance-just activities. As a result, few of these attempts have been successful.

4.3. Maintenance Strategies

Determine the parameters

The initial step of the predictive maintenance is to affirm condition monitoring parameter, affirm parameter measurement method (visual, general instrument measurements and special instruments measurements) as the current and voltage of the electrical equipment, the oil temperature, current and pressure of the crusher, the speed and vibration of the pivot device, and those parameter limits is used as a criterion to monitor.

Detection, monitoring

After determining the parameters, the periodic (eg, weekly, month to month, etc.) or aperiodic, (for example, online random monitoring) approach can be adopted to detect and monitor the process. Additionally, the monitoring methodology and the used instruments is importance, different parameters for different devices and monitoring instrumentation, its monitoring methods are likewise different. When the measured value exceeds the parameter furthest reaches of engineering standards, it is necessary for further examination and diagnosis.

Fault diagnosis

There are many fault diagnosis method, utilizing fault diagnosis technology of spare parts and equipment fault has been diagnosis.

Maintenance work orders

After the results of the diagnosis, the maintenance program has developed, including maintenance personnel, maintenance tooling, maintenance resources, maintenance procedures, and spare parts and so on.

Project Maintenance

As indicated by maintenance program, Project Leader organizes relevant personnel to service and change breakdown status parameter value to the ordinary range. After adjustment or repair equipment, if the equipment have been tested and meet the project the standard range, which can enter a new predictive maintenance cycle.

4.4. Diagnosis and Prognosis of Failures

There is a long history of the technical diagnosis/diagnostics contrasting with technical prognosis/prognostics, which is quite new field of research interest. Prognosis and diagnosis are the key players in service arranging, maintenance and in limiting the down state of the equipment (aerospace is one of the basic area). Ceaseless increase of embedded system calculation performance enables deployment of complex indicative calculations in places, where it was not realistic several years prior. A huge number of data examination are moved from specialized calculation center directly into a monitoring systems and enable us to evaluate conditions in real-time. Diagnosis focuses on detection, detachment and identifies failure when they happen contrasting with prognosis, which focuses to predict failure before they happen. It means that technical prognosis could be understood as an extending/complementary element of technical diagnosis.

We are able to determine the current state as well as we are able to predict future state with some relevance and level of likelihood based on the element and component degradation by utilizing diagnosis and prognosis. The primary objective of the technical prognosis is to make end of life (EOL) and remaining useful life (RUL) predictions that enable timely maintenance decision to be made . Prognostics ought to be performed at the component of sub-component level and ought to involve predicting the time progression of a specific failure mode from its incipience to the time of components failure (Eisenhardt KM, 1989).

There is a noteworthy number of prognostic approaches yet the scientific categorization isn't clearly defined and consensually agreed yet. Most regular characterization parts prognostics approaches to free principle gatherings: model-based prognosis, data-driven prognosis and experience-based prognosis.

A data-driven methodology use the usually observed operating data (currents, voltages, calorimetric data power, vibration and acoustic signs, temperature, pressure, oil debris,) to track, approximate and forecast the system degradation behavior. Measured info/output data is the real source for getting a better understanding of the system degradation behavior. The data-driven (DD) approaches rely on presumption that the statistical data are relatively unchanged unless a failure happens in the system. The regular cause varieties are entirely due to uncertainties and random noise and special cause varieties (e.g. due to degradations) represent not attributed to regular cause. The data driven prognosis is based on statistical and learning techniques from the theory of pattern recognition.

These range from multivariate statistical methods (static and dynamic principle component, linear and quadratic discriminant, fractional least squares and standard variate investigation) to discovery methods based on counterfeit neural networks (probabilistic neural networks, multi-layer perceptrons, outspread basis capacities), graphical models (Bayesian networks, hidden Markov model), self-sorting out feature maps, flag examination (filters, auto-regressive models, FFT etc.), decisions trees and fluffy rule based systems. The vast majority of the work in data-driven prognostics has been applied in areas of auxiliary health management (Eisenhardt KM, 1989). A large number of those systems use vibration sensors to monitor the health of turning machinery, for example, helicopter gearboxes. Some systems monitor the exhaust gases or the oil stream from the engine for defilement that could indicate a fault and it evaluation.

Model-based approaches or supposed material science based are applicable, when relatively accurate mathematical model could be developed from first principle of system's failure modes. Models could be classified as a qualitative or quantitative. The quantitative model represents mathematical and practical relationship between the data sources and yields of a system, while the qualitative models represents these relationships in terms of qualitative capacities centered on different units in the system. Model based approaches are based on investigative redundancy. A process contains scientific redundancy if an info or yield can be calculated by utilizing just other sources of info or yields In the simplest case, the investigative redundancy is utilized by contrasting the yields from the real process and yields from a process model, which is fed by the same contributions as the real process.

Inconsistencies between the model and the real process are represented as a residual. In case of no fault the residual ought to be close to zero (considering the model exactness, flag noise etc.) and on account of a fault the residual ought to be fundamentally non zero on the off chance that it is sensitive to that specific fault. A number of residuals are used and they are made sensitive to different faults to achieve the fault detachment . There is limited number of real application in this area and it could be considered as a most complex and accurate methodology. Use cases defining model based methodology has been created for suspension system in , where simple state-space model was defined and degradation measurement was involved as a gradually evolving feature. Degradation measurements were connected to potential split in suspension system and based on the system stack (Palmer-Miner Law). Reenactments demonstrating this methodology were performed.

Experience/Probability Based Methods

These methods have the longest history, contrasting with other previous approaches does not require too much detailed data and utilize different sorts of likelihood circulation capacities - PDFs, which were parameterized for individual systems/subsystems/components based on creation parameters, operational data, statistical data from history. The most generally used dispersion capacities are ordinary, Weibull and exponential appropriation. A commonplace circulation describing the failure rate versus time is called a "bath curve", which was first time published in 1965 and still has its merit. This prognostics methods additionally provides confidence level in which we operate and we can rely on.

This is critical for determining the likelihood and exactness of our estimate. PDF is used in reliability examination. This methodology is still most normal and is very often applied in the electrical business.

5. e-Commerce Logistics

5.1. The e-Commerce Environment

In defining e-commerce, Kalakota and Whinston (1997) focused on four perspectives: online purchasing and selling, technology driven business process, correspondence of data and customer service. However, this definition does not provide adequate spotlight on exchange cost and other aspects of e-commerce (e.g., B2B, B2G, C2C etc.) Thus, lighting up these basic aspects, Frost and Strauss (2013) extends the definition concentrating on purchasing and selling online, computerized value creation, virtual market places and storefronts and new dissemination intermediaries. However, this definition heavily focuses on e-marketing and fails to integrate other essential e-business processes. All things considered, this examination advances a more comprehensive definition of e-commerce in big data environment, which means to achieve both exchange value (i.e., cost investment funds, enhancing profitability and efficiency) and strategic value (i.e., competitive advantages, firm performance) in computerized markets by changing creation, inventory, development, chance, finance, knowledge, relationship and human resource management with the help of analytics driven bits of knowledge (Wixom et al., 2013).

E-commerce firms are one of the fastest gatherings of BDA adopters due to their need to remain on top of their game (Koirala, 2012). Much of the time, e-commerce firms deal with both structured and unstructured data. Whereas structured data focuses on demographic data including name, age, gender, date of birth, address, and preferences, unstructured data includes clicks, likes, joins, tweets, voices, etc. In the BDA environment, the challenge is to deal with the two types of data with the end goal to generate meaningful bits of knowledge to increase conversions. Schroeck et al. (2012) discovered that the definition of big data incorporated different dimensions including: greater scope of data new sorts of data and investigation real-time data non-conventional types of media data new technology-driven data a large volume of data the latest popular expression and social media data. In defining big data, IBM (2012), Johnson (2012a), and Davenport et al. (2012) focused more on the variety of data sources, while other creators, for example, Rouse (2011), Fisher et al. (2012), Havens et al. (2012), and Jacobs (2009) emphasized the storage and investigation requirements of dealing with big data. As defined by IDC (2013), big data focuses on three principle characteristics: the data itself, the analytics of the data, and the presentation of the results of the analytics that permit the creation of business value in terms of new items or services. Overall, the investigation defines big data in terms of five Vs: volume, velocity, variety, veracity, and value (White, 2012). The 'volume' refers to the quantities of big data, which is increasing exponentially. The 'velocity' is the speed of data collection, processing and dissecting in the real time. The 'variety' refers to the different types of data collected in big data environment. The 'veracity' represents the reliability of data sources. And, at long last, the 'value' represents the value-based, strategic and enlightening benefits of big data (Fosso Wamba et al., 2015b; Wixom et al., 2013).

The sheer volume of academic and industry research provides evidence on the importance of big data in numerous useful areas of e-commerce including marketing, HR management, creation and operation, and finance (Agarwal and Weill, 2012; Bose, 2009; Davenport, 2006; Davenport, 2010, 2012; Davenport et al., 2012). In e-commerce, a large measure of customer-related data is available basically when customers 'sign in': these data are of great interest to business decision makers. While the significance of big data in settling on strategic decisions is recognized and understood, there is as yet an absence of consensus on the operational definition of big data analytics (BDA) (Hesse M 2002).

BDA mentioned in previous studies with the end goal to identify their normal themes. For example, Davenport (2006) indicated that BDA refers to the quantitative investigation of big data with a view to settling on business decisions. Likewise, this decision-usefulness aspect of analytics has been the concentration in other studies, for example, those by Davenport and Harris (2007b), Davenport (2010), and Bose (2009). Whereas Davenport and Harris (2007b) explained BDA with the help of mechanisms, for example, statistical investigation and the use of an explanatory and predicting model, Bose (2009, p.156) described BDA as the "gathering of tools" used to extract, interpret data and also predict the outcomes of decisions.

In defining BDA, one stream of research has focused on strategy-led analytics, or analytics that create sustainable value for business. For example, LaValle et al. (2011) explained that the use of business analytics (or the capacity to use big data) for decision influencing must to essentially be connected with the association's strategy. Indeed, strategy-driven analytics has received much attention due to its role in better decision making. Studies have likewise focused on "competitive advantages" and "differentiation", while applying analytics to analyze real-time data (Schroeck et al., 2012). In a comparative vein, Biesdorf et al. (2013) highlighted that it is vital to create an environment where big data, process advancement, frontline tools and people are well-aligned with the end goal to achieve competitive advantages (Hesse M 2002).

5.2. Big data and their distinctive characteristics in the e-commerce environment

The e-commerce landscape today is rising with numerous big data that are being used to solve business problems. As per Kauffman et al. (2012, p.85), the use of big data is skyrocketing in e-commerce "due to the social networking, the internet, mobile telephony and a wide range of new technologies that create and capture data". With the help of cost-effective storage and processing limit, and forefront scientific tools, big data currently enable ecommerce firms to reduce costs and generate benefits with no trouble.

However, analytics that capture big data is different from conventional data in numerous respects. Specifically, attributable to the elements of their unique nature (i.e., voluminous, variety, velocity, and veracity), big data can be easily distinguished from the conventional type of data used in analytics (J. Xingyi, 2008).

5.2.1. Voluminous

With the emergence of web technologies, there is an ever-increasing development in the measure of big data in the e-commerce environment (Beath et al., 2012). These mass quantities of data that ecommerce firms are endeavoring to harness to improve their decision-production process are defined as voluminous (McAfee and Brynjolfsson, 2012). As illustrated by Russom (2011), BDA takes a large volume of data that require a massive measure of storage and entail a large number of records. Truth be told, BDA encompasses large volumes of data (regularly expressed in petabytes and exabytes) that are used by decision makers for settling on strategic decisions. Data collected in the big data environment are often unstructured and can incorporate video, image, or data generated from mobile technology. All things considered, it is unlikely that big data will be clean and free from any errors. While this poses an extra challenge for decision makers to get data ready for use, big data enable real-time decision making for e-commerce firms (Kang et al., 2003). For example, utilizing the massive measure of structured and unstructured data, Amazon developed sophisticated recommendation engines that deliver over 35 percent all things considered, automated customer service systems to ensure superior customer fulfillment and dynamic evaluating systems that modify valuing against competing sites every 15 seconds (Goff et al., 2012).

Correspondingly, Netflix, the online movie retailer, analyzes over 1 billion reviews to determine the customer's movie tastes and inventory conditions (Davenport and Harris, 2007b). Numerous e-commerce firms (e.g., Amazon, eBay, Expedia, Travelocity) use massive volume of social media data (e.g., photos, notes, blog entries, web connections, and news stories) to take advantage of the chance of real time special offers (Manyika et al., 2011). Notwithstanding opportunities, the volume of big data brings challenges, especially integration of big data from different sources and configurations, presenting new "agile" explanatory methods and machine-learning techniques, and increasing the speed of data processing and examination. All things considered, E-commerce firms must have the capacity to embed analytics and streamlining into their operational and decision processes to enhance their speed and effect (Davenport, 2013a).

5.2.2. Variety

The word 'variety' denotes the way that big data originate from numerous sources which can be structured, semi-structured or unstructured (Schroeck et al., 2012). Variety is another basic attribute of big data as they are generated from a wide variety of sources and organizations including text, web, tweet, sound, video, click-stream, log files, etc. (Russom, 2011). This variety of data requires the use of different investigative and predictive models which can enable data about different practical areas to be used. Biesdorf et al. (2013) explained, for example, that the logical model used by e-commerce firms could comprise a variety of customer data, for example, customer profiles and historical data on purchasing behavior regional and seasonal purchasing behaviors enhancing of supply chain operations and, above all, the retrieval of any unstructured data from social media to predict purchasing by item, store, and advertising activities.

For example, Manyika et al. (2011) showed that an e-retailer provided real-time responses in marketing efforts, amending them as and when necessary by leading sentiment investigation. Overall, the variety of big data can possibly increase the value of firms. However, top management commitment in terms of enhancing business processes and defining workflows is very huge all together for the benefits from such data to be realized (Beath et al., 2012)

5.2.3. Velocity

Velocity refers to the frequency of data generation and/or the frequency of data delivery (Russom, 2011). It is essential to understand the velocity of big data which needs to be prioritized and synced into business processes, decision making and improvements in performance (Beulke, 2011). As described by Gentile (2012), the term 'velocity' is the rate of change in big data and how rapidly big data ought to be used in business decisions with the end goal to include value. Actually, given that greater data velocity is assured, data have the potential to open up new opportunities for associations. As appeared by Davenport and Patil (2012), the high velocity of BDA can enable investigators to direct consumer sentiment examination and provide a clear picture about choices of brands and/or items.

To capitalize on this high pace of data, numerous e-commerce firms have used different techniques to increase the value of their business.

For example, Amazon has been able to keep up a steady flow of new items by right-time correspondence with its stakeholders (Davenport, 2006). eBay Inc. has performed thousands of experiments utilizing data velocity with different aspects of its website, which has resulted in better format and website features running from route to the size of its images (Bragge et al., 2012). To utilize the high velocity of data, numerous e-commerce firms currently use sophisticated systems to capture, store, and analyze the data to make real-time decisions and retain their competitive advantages.

5.2.4. Veracity

Another essential attribute of big data relates to the uncertainty associated with certain types of data. These data demand thorough verification, requiring full compliance with quality and security issues. High data quality is a vital requirement of BDA for better predictability in the e-commerce environment (Schroeck et al., 2012). Therefore, verification is necessary to generate authenticated and relevant data, and to have the capability to screen out terrible data (Beulke, 2011). Actually, verification is essential in the data management process because the existence of awful data may hinder management in settling on conscious decisions. Likewise, terrible data would have little relevance in including business value. Beulke (2011) explained that the data technology (IT) units of e-businesses can assume a key role in this regard by setting up an automatic verification system with the goal that the big data used for business decisions have been authenticated and have passed through strict quality compliance procedures.

In this regard, Schroeck et al. (2012) argued for the use of data combination which combines different less reliable data sources with the end goal to create a more precise and worthwhile data point, (for example, social comments affixed to geospatial area data). Ferguson (2012) highlighted that

Montage Analytics has developed a tool which is especially useful for predicting 'dark swans ' in associations, and some other types of hazard that have originated as a result of human manners and inspirations. The reason is that the inherent unpredictability of some data is constantly caused by factors, for example, technology failure, human absence of truthfulness and economic factors (Jong-Ho Shin and Hong-Bae Jun 2015).

5.3. Types of Big Data Used in e-Commerce

E-commerce refers to the online exchanges: selling merchandise and ventures on the internet, either in one exchange (e.g., Amazon, Zappos, eBay, Expedia) or through a continuous exchange (e.g., Netflix, Match.com, LinkedIn etc.) (K. Bakshi 2012) E-commerce firms extending from Amazon to Netflix capture different types of data (e.g., orders, baskets, visits, users, referring joins, keywords, inventories perusing, social data), which can be comprehensively classified into four categories: (an) exchange or business action data (b) click-stream data (c) video data and (d) voice data (see Appendix 3).

In e-commerce, data are the key to track consumer shopping behavior to personalize offers, which are collected over time utilizing consumer perusing and value-based focuses. This section discusses different types of big data alongside their suggestions for e-commerce.

5.3.1. Transaction or business activity data

Transaction or business movement data evolve as a result of exchanges between the customer and friends over time. These data are structured in nature and originate from numerous sources going from customer relationship programs (e.g., customer profiles maintained by the organization, the occurrence of customer grumblings) through to sales exchanges. A recent report by Chandrasekaran et al. (2013) provided the example of an e-retailer that analyzes data from its unwaveringness program (i.e., its Clubcard dedication program), entailing 1.6 billion data focuses, 10 million customers, 50,000 stock keeping units (SKUs), and 700 stores, which has resulted in the exhaustive coordination of big data with consumer bits of knowledge. In the context of e-retailing, Kiron et al. (2014b) reported that StyleSeek, the online recommendation engine in the US, makes massive revenue by examining customers' tastes and preferences and driving consumers toward its retail partners with the help a sophisticated analytics stage. Overall, it is evident that e-retailers can derive numerous benefits over the value chain utilizing exchange data.

5.3.2. Click-stream data

Click-stream data originate from the web and online advertisements, and from social media content, for example, the tweets, web journals, Facebook divider postings, etc. of e-commerce businesses. In today's connected environment, social media and online advertisements assume a key role in the progressing limited time strategy of firms, for example, the use of click-stream data that are very vital for management in making informed, strategic, and strategic decisions. Earlier studies have discovered that numerous e-commerce firms worldwide (e.g., Amazon, eBay, Zappos, Alibaba etc.) rely on click-stream data in their efforts to capture data. Click-stream data can be applied to predict customer preferences and tastes. As highlighted by Davenport and Harris (2007a), Netflix, a world-acclaimed internet TV network, captures and analyzes more than one billion web data related to reviews of movies that are liked, loved, hated, etc. with the end goal to understand customers' tastes (M. Jeseke, M. Grüner, F. Wieß 2013).

Another recent examination by Davenport et al. (2012) reported that credit card firms, through relying on website and call center data, look after databases (named as ready-to-market) to offer customer-tailored items inside milliseconds and likewise optimize offers by following up responses from customers. Some companies use such databases not exclusively to approach customers yet additionally to offer online services. For example, Biesdorf et al. (2013) explained that by examining web data, e-retailers receive a red banner alert when the prices of their competitors' items are below their very own price level. Therefore, retailers can change their prices to remain competitive.

5.3.3. Video data

Video data are live data that come from catching live images. Currently, e-commerce firms are keen to use not just either click-stream data or exchange data be that as it may, in relationship with image investigation software, they tend to likewise capture video data. As indicated by Schroeck et al. (2012), ecommerce firms have the necessary competencies to analyze extremely unstructured data, for example, video or voice data. These data have the potential to include value for e-commerce firms. For example, Ramaswamy (2013) reported that Netflix uses video data to predict viewing propensities and evaluate the nature of experiences. Moreover, the representation and demand analytics tool based on the type of movie utilization help Netflix understand preferences, which led them to achieve success in their "House of Cards" program in the US. Along these lines, the use of video data is essential for firms in settling on better decisions than their competitors (N. Marz and J. Warren 2012).

5.3.4. Voice Data

Another type of data attached to the big data family is voice data, that is, data regularly starting from phone calls, call centers, or customer service. As evidenced in recent research, voice data are advantageous for breaking down consumer-purchasing behavior or targeting new customers.

As described by Davenport et al. (2012), credit card companies, for example American Express, use and track data related to call center activities so personalized offers can be given in milliseconds. In Schroeck et al's. (2012) survey, e-commerce firms were found to use advanced capabilities to analyze text and transcripts converted from call center conversations. Moreover, numerous nuances of language, for example, sentiment, slang and intentions, can be read and recognized by means of BDA in the context of e-commerce (W. Hofman 2011).

Since the nature and type of big data are unique and originating from different networks of advanced stages, there is probability of new theory enquiring new problems. The data economy additionally indicates that big data are "relational" and "networked", which necessitate new developments in IT capability and calculations, system and data quality, security and ethical ramifications, strategic alignment and corporate culture.

5.4. Organizational and Business Value of Big Data for e-Commerce Firms

The ultimate challenge of BDA is to generate business value from this explosion of big data (W. Hofman 2011) The term 'value' in the context of big data implies the generation of economically commendable bits of knowledge and/or benefits by breaking down big data through extraction and change.

Aligned with Wixom et al. (2013), we define business value of BDA as the value-based, instructive and strategic benefits for the e-commerce firms. Whereas value-based value focuses on enhancing efficiency and cutting costs, instructive value sheds light on real time decision making and strategic value deals with increasing competitive advantages.

For example, by injecting analytics into e-commerce, managers could derive overall business value by serving customer needs (79%) creating new items and services (70%) expanding into new markets (72%) and increasing sales and revenue (76%) (Columbus, 2014).

Amazon, the online retailer monster, is an exemplary example of enhancing business value and firm performance utilizing big data. Indeed, the firm could generate about 30% of its sales through analytics (e.g., through its recommendation engine) (The Economist, 2011). Essentially, Kiron et al. (2012b) reported that Match.com could earn over half increase in revenue in the previous two years while the organization subscriber base for its core business reached 1.8 million. The IBM case think about (IBM, 2012) illustrated that greater data sharing and analytics could improve patient outcomes. For example, Premier Healthcare Alliance could reduce expenditure by US$2.85 billion. Schroeck et al. (2012) discovered that Automercados Plaza's grocery affix could earn a nearly 30% rise in revenue and a total of US$7 million increase in benefit each year by implementing data integration all through the association.

Furthermore, the organization avoided losses on over 30% of its items by scheduling price reductions to sell perishable items on time. Notwithstanding including value for business in money related terms, the use of big data can include benefit in non-monetary parameters, for example, customer fulfillment, customer retention, or enhancing business processes. As presented by Davenport (2006), United Parcel Service (UPS) examines usage patterns and grumblings data to accurately predict customer defections. This process has resulted in a critical increase in customer retention for the firm. In the comparable soul, LaValle et al. (2011) reported that an online automobile organization could develop accurate customer retention strategies by creating a customer sample from big data followed by applying diagnostic calculations to forecast weakening probabilities, coupled with identifying in danger customers. This retention strategy consequently has opened up prospects for the firm to yield hundreds of a huge number of dollars merely from a single brand. Because e-commerce firms have opportunities to interact in real time with customers more frequently than firms that don't have an e-commerce stage, they need to use big data for different purposes.

6. Methodology

Secondary analysis (SA) is a term used to describe different expository practices which make use of pre-existing data either to investigate new research questions or to re-examine essential examination questions for purposes of corroboration. Data used in this context are thereby transformed into 'secondary data'. This data, which can be quantitative or qualitative or a mixture of the two, may have been collected in the context of a research think about (hereafter referred to as 'data sets') or for other purposes (referred to as 'naturalistic' data, for example, diaries and administrative records). In spite of the fact that meta-examination and literature reviews involve the use of crude and/or published data from existing studies, these methodologies can be distinguished from SA in that they are concerned with synthesizing research discoveries, rather than examining new research questions or verifying results from individual studies.

7. Conclusions

More and more companies implement a "data-driven" mindset inside their decision-makers: an investigation by the Council of Supply Chain Management Professionals demonstrates that 93% of shippers and 98% of outsider logistics firms feel like data-driven decision-production is urgent to supply chain activities. 71% of them believe that big data improves quality and performance. These figures indicate how essential it becomes for everyone in the sector, from retail to supply chain.

Other benefits of utilizing big data are the resource utilization streamlining and the improvement of operational efficiency:

Forecast: having experiences on how customer demand will evolve helps to design and anticipate shifts, inventory shortages, and reduces costs.

Inventory Management: related to storing and forecasting, the inventory management is enhanced. RF labels and computer chips make it possible to keep track of inventory and prevent unpredicted stockouts.

Route streamlining: on account of real-time GPS data, weather data, street maintenance data and fleet and personnel schedules integrated into a system taking a gander at historical trends, the most optimized routes and time are selected for delivery. An exemplary delineation is the means by which UPS implemented a "no-left-turn approach", which as indicated by the organization, saved them yearly 10 million gallons of fuel, delivers 350,000 more packages and emits 20,000 tons less of carbon dioxide.

Work management: on account of analytics, staff management is made easy, which is a huge asset for the supply chain where a ton of workforces is needed on a ceaseless basis. Ensuring enough resources are available in anticipated surge periods maintains a strategic distance from overtime and exhaustion.

The move in customer experience initiated by e-commerce has put the ball in the customer's court, and definitely changed the relationship between them inside the retail business. The augmentation of players in a single area has generated a competition that ruins the customer to pull in and keep him/her: very low delivering costs, flexible return policies, regular limits, etc. Big data was of great help in the management of these policies.

Another aspect where big data helped retailers with, is the increased knowledge about the customers themselves: website analytics enable to track the audience foundation (age, gender, area, interests, etc.) and their behavior. That takes into consideration better targeting and segmentation, more effective marketing efforts, as you get a better understanding of your customers that are in return more steadfast. Examining consumer patterns gives you opportunity to get better and improvement of retail strategies.

Automation is everywhere in today's business world, increasing speed, precision and efficiency of processes in capacities like marketing, recruitment or finances. It frees up a great deal of time for your teams to center around less repetitive and time-expending undertakings, to use their added-value on work that really matters.

Likewise, the logistics sector is affected, and soon enough it will be entirely automated.

A precursor in the matter is, of course, Amazon, the e-commerce monster that invests in pioneering technologies and that is permanently upsetting the retailing and online retailing landscape – and now assaulting new markets like grocery stores. In the Amazon warehouses, the little orange KIVA robots snatch the items from the shelves and convey them to the employee that packs them. Amazon has additionally tried drone delivery for people living less than 30 minutes from an Amazon center.

Seeing how Uber and other Tesla companies are beginning to test-run self-driving vehicles, envisioning that the whole supply chain could be automated is only one step away. It does not mean that robots will take everything over, as there will dependably be the need of the reassuring human interaction, however automation is definitely going to thrive in the coming years and decades.

Big data analytics (BDA) has emerged as the new frontier of advancement and competition in the wide spectrum of the e-commerce landscape due to the challenges and opportunities created by the data revolution. Big data analytics (BDA) increasingly provides value to e-commerce firms by utilizing the elements of people, processes, and technologies to change data into bits of knowledge for strong decision making and answers for business problems. This is a comprehensive process which deals with data, sources, aptitudes, and systems with the end goal to create a competitive advantage. Leading e-commerce firms, for example, Google, Amazon, and Facebook have already embraced BDA and experienced enormous development. Through its systematic review and creation of scientific categorization of the key aspects of BDA, this examination presents a useful beginning stage for the use of BDA in emerging e-commerce research. The examination presents a methodology for encapsulating all the best practices that construct and shape BDA capabilities.

Also, the investigation reflects that once BDA and its scope are well defined, distinctive characteristics and types of big data are well understood, and challenges are properly addressed, the BDA application will maximize business value through encouraging the pervasive usage and speedy delivery of bits of knowledge crosswise over associations.

8. References

A Starr, R Willetts, P Hannah, W Hu, D Banjevic, and AKS Jardine. Data fusion applications in intelligent condition monitoring. In IFRIM Conference, pages 6–8, 2002.

A. Katal, M. Wazid, R. H. Goudar, “big data: Issues, Challenges, Tools and Good Practices”, IEEE Sixth International Conference on Contemporary Computing (IC3), pp. 404-409 , 08.2013

Andrew K.S. Jardine, Daming Lin, and Dragan Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing, 20(7):1483 – 1510, 2006.

C. Sooriaarachchi, T. Gunawardena, B. Kulasuriya, T. Dayaratne, “A study into the capabilities of NoSQL databases in handling a highly heterogeneous tree”, IEEE 6th International Conference on Information and Automation for Sustainability (ICIAfS), ISBN 978-1-4673-1976-8 , pp. 106-111 ,2012

D. Pfisterer, et al., “SPITFIRE: Toward a Semantic Web of Things”, Communications Magazine, IEEE (Volume: 49, Issue: 11), pp. 40-48, 2009

Eisenhardt KM, (1989) Building theories from case study research. Academy of Management Review. 14(4): 532-550.

Elsayed A. Elsayed. Reliability engineering. J. Wiley, Hoboken, N.J, 2nd;2. aufl.;2; edition, 2012.

Gersick C (1998) Time and transition in work teams: Toward a new model of group development Academy of Management Journal. 31: 9-41.

Hesse M (2002) Shipping news: the implications of electronic commerce for logistics and freight transport. Resource Conservation & Recycling. 36(3): 211-240.

J. Xingyi, "Efficient Complex Event Processing over RFID Data Stream", Seventh IEEE/ACIS International Conference on Computer and Information Science, ISBN 978-0-7695-3131-1, pp. 75-81, 2008

Jong-Ho Shin and Hong-Bae Jun. On condition based maintenance policy. Journal of Computational Design and Engineering, 2(2):119–127, 2015.

K. Bakshi, "Considerations for big data: architecture and Approach ", Aerospace Conference, 2012 IEEE, ISBN 978-1- 4577-0556-4, pp 1 – 7, 2012

L. Zhang, X. Li, and J. Yu. A review of fault prognostics in condition based maintenance. Proc. of SPIE, 6357:635–752, nov 2006.

M. Jeseke, M. Grüner, F. Wieß, "BIG DATA IN LOGISTICS: A DHL perspective on how to move beyond the hype", DHL Customer Solutions & Innovation, 12.2013

N. Marz and J. Warren, "Big data. Principles and best practices of scalable realtime data systems". Manning Publications, MEAP Edition, Manning Early Access Program big data version 7, 2012.

Otilia Elena Dragomir, Rafael Gouriveau, Florin Dragomir, Eugenia Minca, and Noureddine Zerhouni. Review of prognostic problem in condition-based maintenance. In Control Conference (ECC), 2009 European, pages 1587–1592. IEEE, 2009.

Phil Blunsom. Hidden markov models. Lecture notes, August, 15:18–19, 2004. [10] Nitish Srivastava, Geoffrey E. Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15(1):1929–1958, 2014.

R. Kothamasu, S. H. Huang, and W. H. VerDuin. System health monitoring and prognostics — a review of current paradigms and practices. The International Journal of Advanced Manufacturing Technology, 28(9):1012–1024, 2006.

S. Ghemawat, H. Gobioff, S.T. Leung, "The Google File System", ACM SIGOPS Operating Systems Review, ACM, pp. 29-43, 08.2003

S. Robak, B. Franczyk, M. Robak, "Applying big data and linked data concepts in supply chains management," 2013 Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 1215 – 1221, 201

T. Berners-Lee, et al., "W3C Semantic Web Activity" [Online]. Available: http://www.w3.org/2001/sw, 2001 [7] S. K. Rajapaksha, et al., "Internal Structure and Semantic Web Link Structure Based Ontology Ranking", ICIAFS 2008. 4th International Conference on Information and Automation for Sustainability, pp. 86-90, 2008

W. Hofman, "supply chain visibility with Linked Open Data for supply chain Risk Analysis", Workshop on IT Innovations Enabling Seamless and Secure supply chains, pp. 20-31, 08.2011

Weltevreden JWJ (2008) B2E e-commerce logistics: the rise of collection and delivery points in the Netherlands. International Journal of Retail and Distribution Management. 36(8:) 638-660.

Xing Y, Grant DB (2006) Developing a framework for measuring physical distribution service quality of multi-channel and “pure Player” internet retailers. International Journal of Retail and Distribution Management 34(4-5): 278-289/

Yin RK (2003) Case Study Research: Design and Methods. 3nd edition, Thousand Oaks, Sage Publications.

Ying Peng, Ming Dong, and Ming Jian Zuo. Current status of machine prognostics in condition-based maintenance: a review. The International Journal of Advanced Manufacturing Technology, 50(1):297–313, 2010.