TOWARDS SUSTAINABLE INDUSTRIALIZATION USING BIG DATA AND INTERNET OF THINGS

Dr. V. Suma,
Professor
Department of Information Science and Engineering,
Dayananda Sagar College of Engineering,
Bangalore, India.
Email: sumavdsce@gmail.com

Abstract: The developing nations requires to have sustainable developments in the industries to have new product development and continuous improvement in the social and economic levels paving way for the customer satisfaction. The Big data and the Internet of things are innovative technologies that have made their separate stand in the industrial sector of manufacturing and providing competitive edge to the companies. So the combining of two technologies would multiply anticipation and would create way for the enhanced productivity by satisfying the global needs of the business. Further, the internet of thing along with the big data analytics is also expected to provide with the enhanced security provision, quality of process and products with reduced production down time and financial risk. So the paper proposes the big data and IoT role in the industrial sector and evaluates its performance by analyzing the impacts of the big data and the internet of things in the industry to have sustainability in production process. The proposed framework also analyses the potential applications and the key advantages offered by the integration of the internet of things and the big data.

Keywords: Internet of things, Big-data, Sustainability, Industries, Production, Quality of life improvement.

1. Introduction

The emerging technological growth, IoT, and social media causes swift changes in the modern world, producing humongous data flow. This large data flow which was once considered to be a junk is now gathered to extract useful information to pave way to the new developments, creating new solutions for raising the social and economic levels, to ensure an enhancement in the quality of life. The processing of the large amount data for the extraction of the value using the conventional methods of data processing and application software has become impossible due to the complexity of the data gathered. So the emergence of the big data enables the ways in analyzing and systematically extracting values from the information's that are highly complex and conceived with many cases. The insights provided by the big data to achieve development in a field has made it to gain more popular among the industries and academia. The various Industrial organization such as banking, government, manufacturing, education, retail,

I-SMAC

and healthcare benefits when armed with the insights provided by the big data in financial assistance, solving problems easily, managing utilities cleverly, improving patient care and developing customer relationship respectively leading to the better decisions and strategic moves for the sustainability. The opportunities of including big data also provides with the better intelligence and the value- creation in the industries. As mentioned above the IoT is the major source of data flow that cannot be comprehended easily, as more and more devices are getting connected over the internet day by day and big data analytics becomes a crucial part in the Iot as it helps in the analyzing of the enormous data generation in the real-time and afford a better decision making. Although both technologies have evolved independently they have become interrelated over the period of time. The big data tools poses the capability of handling large volumes of data generated from IoT devices. The internet of things service that are unique create suitable opportunities to enhance the customer value using the big data analytics. Consider the example of the enormous data generation of the sensors. These humongous data gathered are analyzed in real-time using the big-data analysis to derive values from the data, by observing its patterns and trends which may be informative in providing decisions for continuous improvements in the operations or the functioning.

So the Iot and the big data are interrelated, in which the Iot has become the driving force of the big data by causing huge amount of data gathering from several different sources in our day to day life to enable new perspective by aligning the technologies in the best possible way.

The IoT and big-data go hand in hand playing an important role in the industrial sectors. So the paper proposes the performance analysis of the combining the two technologies to have sustainability in the industrial production.

The remaining of the paper is arranged as 2 Related Works, on the big data its challenges and the effective analyzing in extract the values from the information. 3 Proposed work, Gives details about the impact of combining the IOT and Big data for industrial sustainability 4 Performance analysis, showing sustainability the manufacturing, logistics and supply chain 5 Conclusion

2.Related works

De Mauro et al [1] the paper gives the elaborate description on the trending big data that is used in processing the each data gathered in the day to day life through different means, to extract useful information, in aid to the developing society. Jones et al [2] the paper proposes the process of framing a concept for the development of the individual, a group or community or a firm, by gathering the data and differentiating them as fundamental data as they are collected and practical data as they are in use and checks for their effect on data-processing. Mikalef et al [3] the past work on the IT business research are gathered and the resource based and the dynamic capability view of

I-SMAC

the concern is framed using the big data analytics to provide with the mechanisms in strategically leveraging the entailed organizational changes for the improvement of the industry by creating awareness on the different business values, different business elements and the interdependencies between them. Tonidandel et al [4] the paper scopes in initiating the interest in the big data analytics by promoting the biggest advantages and the disadvantages proffered by the big data in the organization science, so as to help in the future research involving the big data analytics. Sun et al [5] the research and development in the business analytics is provided by the big data analytics in enhancing the business intelligence. This also includes the big data science and big data computing for the designing of a big data analytic service oriented architecture to enhance the temporal, expected and the relative characteristics of intelligence in the business organization. Amado, et al [6] the survey analyzing the text mining in a semi-automated approach to find out the trends in the marketing by including the analysis by focusing on the big data, marketing, country continent, product and sector. The research showed that the applications require more development to thrive into the marketing. Storey et al [7] the review on the five V's of the big data is presented including the NoSQL database that helps in accommodating the big data needs and initiatives for effectively using the enormous data collected in the decision making and knowledge discovery. The enhancing's for the conceptual model is also suggested by analyzing it with respect to the big data. Akter et al [8] the study uses the big data analytics capability model categorized into three sections as management, technology and talent capability to analyze the capabilities of the big data in enhancing the performance of the firm. Akter et al [9] the study on the big data aspects, characteristics, types, business value and challenges in the E-commerce is presented and the overall study results, providing with the deeper insights along the cross-cutting analytics application in the E-commerce. Sivarajah et al [10] the paper provides the systematic literature review by the observation and the understanding of the past trends in the big data analytics to identify the potentials, contributions, limitations and implications to aid the future research and the patterns. Hu et al [11] the survey on the big data provides the new finding in the big energy data analytics and security for energy planning, efficient energy generation and distribution. Lv et al [12] the challenges in the big data analytics in the traditional research field is analyzed in the paper, this analysis provides the solutions on solving the challenges with the big data including various types of network. Zhou, et al [13] the smart planning, key technologies, opportunities, challenges for the development of future trend industry with manufacturing process that is intelligent using the big data analytics in the formulation of system standardization and efficient management of the industry. Kiel et al [14] the paper is study that is related to the economic, ecology, social benefits and challenges for creating a database for the sustainable industrial development. The study concludes with the findings that the internet of things require a triple bottom line extension in technical integration, data and information, and public context for the sustainable industrial value creation. Sun et al [15] the ability of the internet of things and the potential of the big data is analyzed in this paper to provide that the conjunction of the IOT and big data causes considerable changes in the community providing the ubiquitous networking and a real time control respectively. Kwiatkowski et al [16] the paper uses the big data and the internet of things to introduce more and more new solutions for the industrial development in the logistics and supply chain management. Koo, et al [17] the paper proposes the myriad information collection by IOT through wireless sensor networks for the big data analytics in the downstream and upstream data collection on water supply to manage the water usage and achieve higher level of

I-SMAC

sustainability in the water supply. Maksimovic et al [18] the use of the data collection and analysis with big data and IOT for the health industry is proposed for developing a sustainable health care industry and identified the priorities of IOT in the area of healthcare. Channe et al [19] the industrialization in the smart agriculture is done using the mobile app which is aided with the big data, cloud computing, mobile computing and so on. Mourtzis et al [20] the IOT application is designed and developed for the real industrial data generation leading to industrial big data.

3. Proposed Work

The modernization of industries with the machine tools enabled with the embedded sensors, cloud based solutions and the business management, caused an rapid increase in the generation of the data , these data generated which were treated as junk , was later recognized to have values which and this created the necessity in the processing of the data. It is reported that the industrial data generation which has reached the volume of the 100 Exabyte's and would increase further in the coming years. So there arises the need for the knowledge provisioning to the machine tools for the purpose data capturing and analysis. This is where the IoT and the big-data come into existence in the industrial sectors.

The internet of things is the system of connected tangible commodities through internet. The IoT enables the thing or the person or any device that has a hidden identity to generate and collect information through the internet without manual intervention but with the help of the embedded technology. In industries the IoT plays a significant role by connecting machine and the resources to the internet which ends up with the transferring of the machine into automated machines with the ability of capturing the heterogeneous amount of data from diverse sources, the data captured cannot be simply stored as there are abundant amount of information gathered through IOT and they remain useful only when it is set in action. These enormous data set collection which might be structured, unstructured or semi-structured is processed and analyzed by the means of big-data to provide insights for the latest development in the business to uphold the industrial competition.

3.1. Big-Data in Internet-of-things

The big-data analytics play a significant role in the analyzing the data generated through the internet of things connected commodities to extend support in having an improved decision making. So the big-data for the Internet of things processing all the gathered data's and storing them using different storage technologies. The analysis of the generated data through IOT is followed as per the step given below

I-SMAC

STEP1: Initially the data collected by the IOT connected device, which might be structured, unstructured or semi-structured is gathered by the big-data system depending on the volume, variety and velocity of the data generated.

STEP2: The Generated data is reserved into large data files in the distributed and shared database (Big-data system)

STEP3: Analyses the data using the Hadoop Map Reduce and Machine Learning.

STEP4: Generates the final output data value as the result of the data analyses.

So the data gathered over internet through the IOT is in the requirement of the very speedy analysis including more number of queries to gain the rapid insights from the data gathered to have quick decision making. Because of this the big-data becomes an entailment in the IoT and termed as the fuel enabling the internet to run as shown in the Fig 1
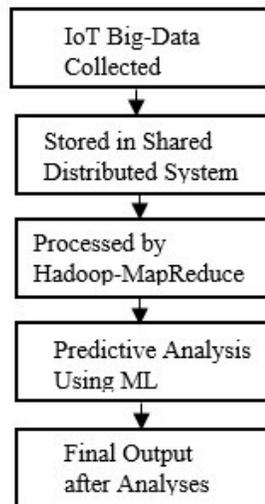

Fig 1 Proposed Flow Diagram

3.2.Interdependency of Big-Data and Internet-of-Things

The Internet of things and the Big-data not only are interdependent they also impact each other as shown in Fig 2. As the IOT flourishes it places increasing demands on organization regarding the big data capabilities. Consider the growing amount of data generation and also consider the limits of the traditional devices used in the data storage provisioning. This cause the demand for the more innovative and the additional storage devices. To handle the increasing workloads resulting in the refurbishing of the organizations infrastructure for the big data storage.
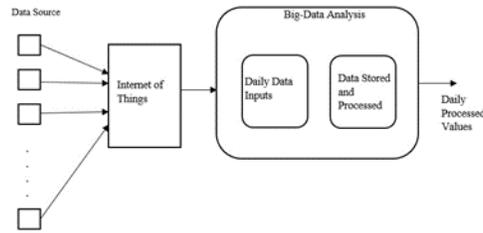
28

Fig 2 IoT Big-Data Interdependence

3.3. The Impact of IoT and Big-Data in Sustainable Industrial Process

The IoT and Big-data plays a significant role in the development of the sustainable production and the distribution of the products. The industries benefit by analyzing the huge amount of the data generation using the internet-of-things and the big-data. They provide with the advanced decisions in managing and identifying the enhanced ways in improving the business and assisting the business and other organization to achieve a detail understanding of data to make efficient and well-informed decisions in all its sectors. This paper particularly concentrates on attaining the production process sustainability using the big-data and the internet of things as they production of the product plays the significant role in the sustainable industrialization.

The Industries are towards sustainability due to the scarcity in the availability of resources, awareness among the consumers regarding the green process and the products, regulations of the environment and cost savings. The sustainable industries are to produce and distribute the products with resource utilization that is flawless and by removing the preventable waste and the toxic products and also has to concentrate on the expense invested, the return to be gained on the investment, technical, information limitation standards and the decision support systems to have a decrease in the energy consumption, environmental cost and business waste.
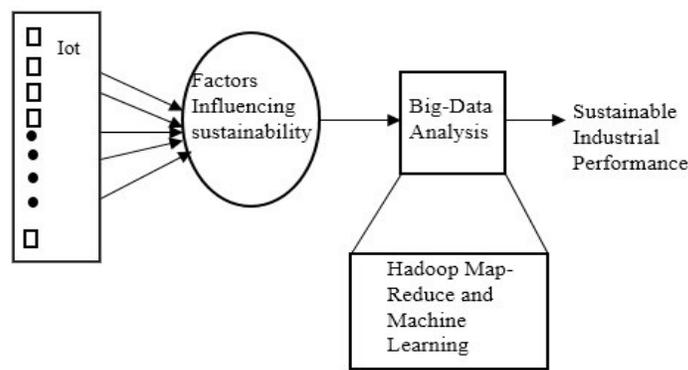

Fig. 3 IoT and BD in Proposed sustainable Industrial Process

29

I-SMAC

The Big data and the Internet of things as shown in Fig 3 pave way for the sustainable industrialization keeping in mind the aspects of social, economic and environmental systems. The IoT and Big-data provide with efficient data capturing and the predictive analysis increases the overall reliability as well as the firm performance. Considering the product development with energy efficiency, resource conservation, safe and skill enhancing working conditions, with reduce in the wastage production, with the aim of meeting the needs of the future generations, and reducing the cost and the execution time, the proposed paper frames a model with factors such as the administration style, supplier convergence, internal business process, quality management, customer convergence, green lean systems, improved quality of performance and government policies.

The data captured using the IoT is based on the factors influencing the sustainability is provided to the big-data analysis. The data collected is initially mapped with number of iterations over the data , aggregated and reduced using the Hadoop-Map-reduce, and followed with the machine learning model that is trained to learn the sub factors of the factors influencing the sustainable industrialization. The factors and the sub-factors influencing the sustainable industrialization are listed in the table.1 below.

	Factors	Sub-Factors
Sustainable Industrialization	Administration Style	Top management commitment, Training and Education for Awareness ,Role of middle-level management, appreciation of employees, Employee involvement, Well- defined sustainability policy Discretion of tasks
	Supplier Convergence	Involvement of supplier, Sharing of sustainability targets with vendors and suppliers, Supplier as a green partner or not, Information sharing structure with the supplier, Cooperation of supplier.
	Internal Business process	Written objective and policy for Sustainability, Routinization Information sharing (IS)
	Quality Management	Cost to quality, Total quality management, Total productive maintenance, Data quality, Quality of Service and Experience
	Customer Convergence	Customer feedback for eco-design, Customer cooperation for cleaner production, Customer cooperation for green purchasing, Information sharing structure with customer
	Green Lean System	Tools and techniques of lean practices, Energy efficient technologies, Green manufacturing, Reuse, recycle, and remanufacture, Reconfigurable Manufacturing Systems
	Government Policies	Support of state/central government, Regulatory pressure to promote sustainable practices, Regional environment to implement sustainable practices, Policy about security and privacy if data
	Improved Quality Performance	Environment technology, Air pollution control, carbon footprint, Eco-packaging, Recycling efficiency, Environment cost, Responsiveness cost, Supply chain cost, Customer satisfaction, Reduction in solid and water waste, Improvement in organizational relevant knowledge

Table 1. Factors and Sub-Factors of Sustainable Industrialization

I-SMAC

3.4.Hadoop Map-Reduce and Machine learning

The Hadoop Map-Reduce (H-MR) is the popular processing engine for the big-data, the Hadoop Map-Reduce Contains two user defined functions map and reduce. The input of the H-MR uses a pair of key values (m, r), the map function is called for each data collected to produce the zero or the more intermediate values from the data. The values are framed into a group and calls for the reduce functions for each group to produce with the more aggregate results. The H-MR process the value for the sub-factors mentioned in the table 1. The algorithm given below shows the steps involved in the optimal MR

```
Input: D (data in HDFS), min_sup (minimum support threshold),
Output: d, frequent itemsets Method:
m1 =find_frequt_1-itemsets (D);
for (r1=2; r −1 != ∅;r1++){
C k =candidate_gen(r −1 ,min_sup);
for each row of data t∈D{//scan D for counts
C t = map();//get the subset of the candidate itemsets
}
r =Reduce();// get the subset of the frequent itemsets
}
return d=r ;
```

Fig 4. Algorithm for H-MR

The values generated for the processed data are given as the training set in the hidden layer for the ANN model developed with input begin defined using the factors influencing the sustainability and the output attained with the key features such as energy efficiency, resource conservation, safe and skill enhancing working conditions, with reduce in the wastage production, with the aim of meeting the needs of the future generations, and reducing the cost and the execution time, enhancing the sustainability in the industrial performance. The Hadoop ensures the advantages of allowing a non-expert user to easily run analytical tasks over big-data, giving full control on the input data set process and taking care of the parallelization and failover in the HDFS with only user defined functions on map and reduce. The Ann model as shown in Fig. 5 considers the reduce output of the H-MR for analyzing the influence of the big-data over the sustainable industrialization using the training set of the sub-factors defined based on the factors influencing the sustainable industrialization.
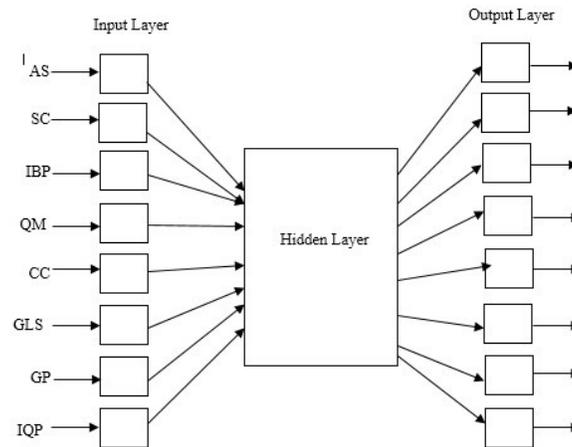
Fig 6 ANN model for BDA

The sustainability in the industrialization using the Iot and the big-data proposed in the paper helps in the proper processing of the factors influencing the sustainability, and further analysis using the learning techniques helps the proposed system of the interdependent big-data and the Iot paving way for the sustainable industrialization.

4.Result Evaluation

The Proposed model enabling the Iot services for the big-data analysis using the machine learning and the H-MR is simulated on the network simulator on the varying number of the industrial cores ranging from 50 to 100 based on the factors and the sub-factors used for processing and the analyses. Under the area of 1000*1000 sq. Units, the usage of the H-MR for the process of data processing ensures advantage of the reduced power consumption, cost, execution time and latency and the proposed process is validated based on the waste production and the productivity increase to ensure the sustainability in the industrial process.

a)Power Consumption
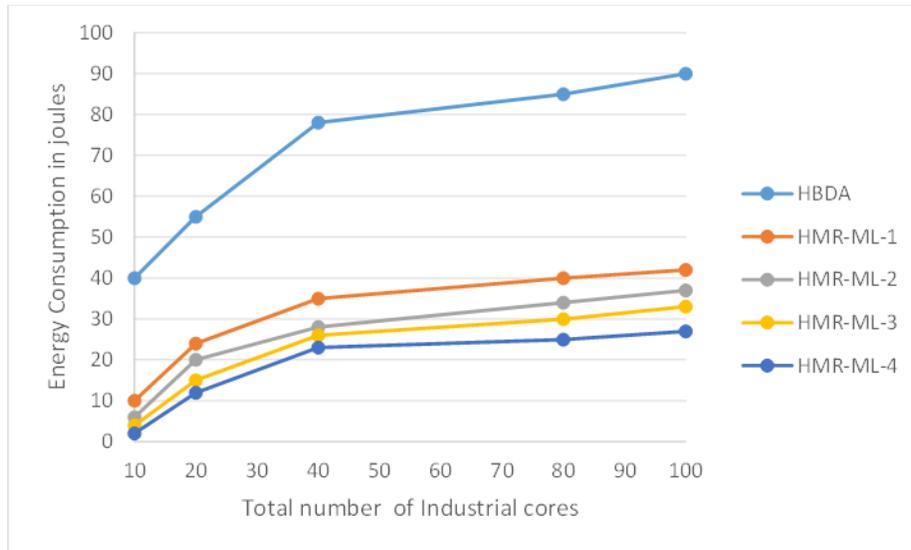
http://irojournals.com/iroismac/

I-SMAC

Fig 7 Power consumption

The power consumption of the proposed model on the analysis of the large data is evaluated to know the efficiency in the sustainable industrialization in terms of the power consumption. The proposed model using the Hadoop and Map reduce, as it allows the user to define only the two functions and the performs all the other process of parallelization and failover by itself and the proper training extended for the analysis of the big-data, helps in the reduction of the overall power consumption for the analyses in the sustainability development causing energy efficiency in the industrial process.

The Fig 7 shows the power consumption minimization, as energy efficiency is one of the main key-features for the sustainable industrialization. The proposed model ensures energy efficiency in industrial process by the proper mapping and reducing and the proper training and testing using the machine learning for enhancing the performance of the industry without human intervention. Which would otherwise cause more time and energy when done manually. The combined Hadoop and the ML shows considerable reduction in the energy compared to the analysis performed just using the Hadoop.

b)Low Latency

The combined Hadoop and ML has caused a considerable reduction in the overall time consumption of the processing and the analysis there by reducing the overall latency in the industrial process, as the human intervention and the manual analyses are avoided due to the automation introduced by the ML- ANN model.
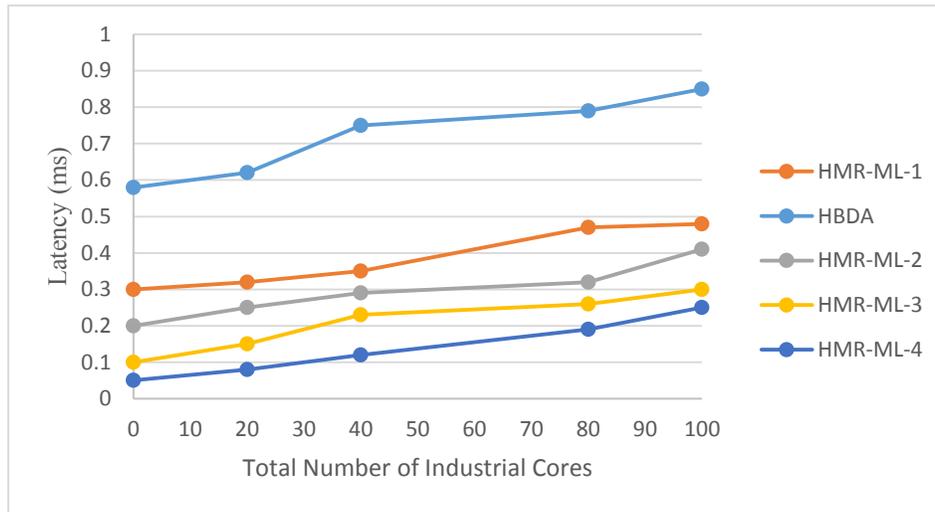

Fig 8 Low latency

The Fig 8 gives the low latency measured for the proposed model that performs the machine learned analysis on the data set that is reduced using the Hadoop Map-Reduce. So this shows a better improvement in the overall system latency by the employment of systems and tools that reduce the latency in the industrialization process based on the analyses performed on the sub factors such as the energy efficient tools, quality service, information sharing and reconfigurable manufacturing.

c)Execution Time

The machine learning analyses made by training set developed based on the sub-factors such as quality management, production maintenance, data quality, quality service, employee involvement, training and awareness and the discretion of the tasks, enables the industries in completing the tasks within time and causes a timely delivery without affecting the quality of the product. As the analysis fully automated the time taken in the analysis of the data are also very much less compared to the analysis made manually.
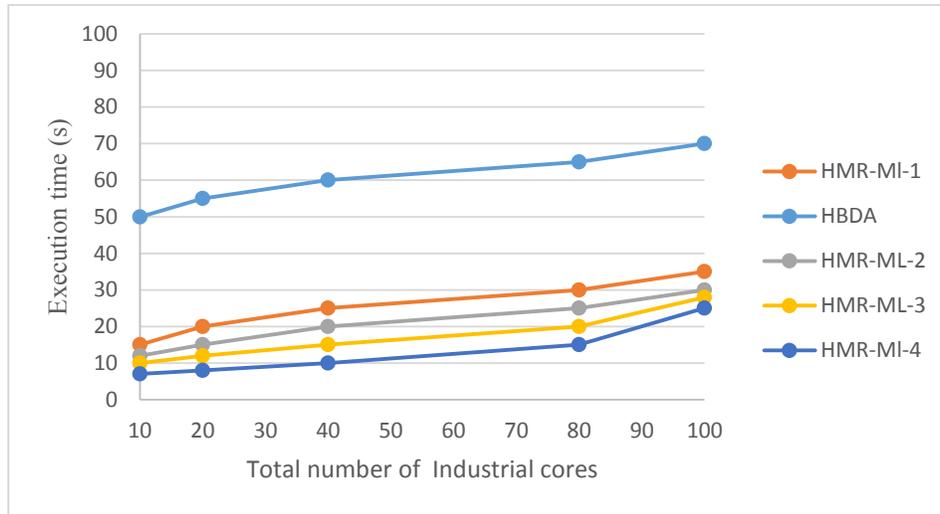
I-SMAC

Fig 9 Execution Time

d)Lower waste production

The data obtained and processed and analyzed based on the sub-factors, such as green manufacturing, ecofriendly designing, reuse, recycle and re-manufacturing, eco packaging, recycling efficiency, air pollution control, carbon foot print, reduction in the solid and the waste, gives valuable data to the industries on how to further reduce the waste and the toxic chemicals. Moreover the valuable information gained by the analysis on the eco-packaging, recycling, air pollution control, reduction in the solid and the toxics gives, the major amount of the industrial waste reduction that has happened in the industrial process so as to make it sustainable and eco-friendly.

The Fig 10 shows the amount of waste products reduced as the result of the data analyzed to improve sustainability in the industrial process thus reducing the harmful side effects caused by the industries. The waste production checked for varying number of the industrial cores has caused a considerable reduction in its production due the analysis leading to the sustainability of the industrial process compared to the existing methods.
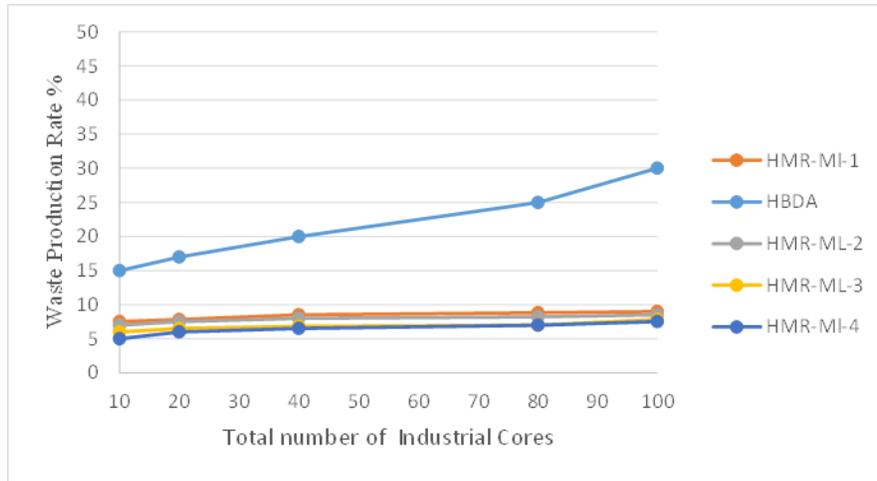
Fig 10 Lower waste Production
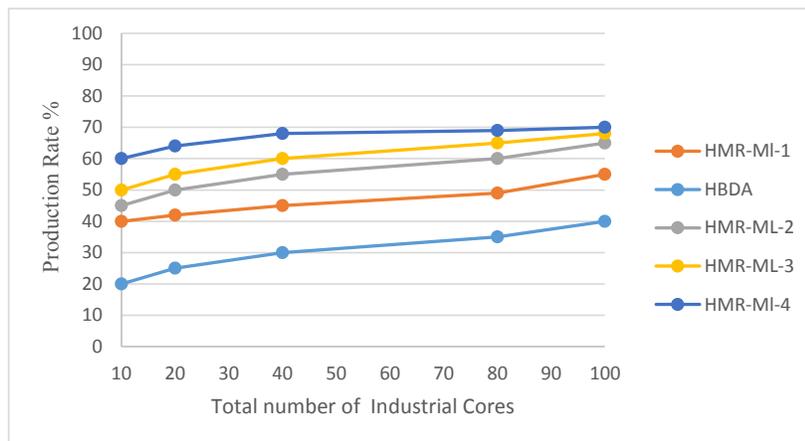
e)Production Increase

Fig 11 Production Increase

The analysis based on the customer feedback, customer co-operation for the green purchasing, eco packaging, and co-operation of the suppliers over the internet and the value extraction from the data collected causes the manufacturer to increase his production based on the customer needs. The valuable data collection made and the processing done over the IoT and the BDA enables the industries to increase their production rate meeting the user

36

I-SMAC

needs, satisfying the customers. And also making the production sustainable meeting the needs of the future generations.

The Fig 11 on the production increase shows the increase in the production rate of the industries due to the sustainable practices adoption. The collected data regarding the customer support for the product and its value extractions has helped lot in the industries to develop their production according to the user needs, and future generation's needs at the reduced toxic and the waste emission. Thus making the industrial process sustainable.

5. Conclusion

The paper proposes the sustainable industrialization using the IoT and the big-data analysis, the proposed model gather the information from the various areas of the industries using the IoT and process them using the Hadoop and Map-reduce and further makes analysis based on the factors that influence the sustainability and their sub-factors, using the machine learning to pave way for the sustainable industrial process in the terms of the energy, resource conservation, product production meeting the present and the future needs , safe and skill enhancing working conditions, low waste production, execution time and cost. The proposed model validated proves to be efficient in terms of the power consumption, latency, execution time, causing reduction in the industrial waste production and increase in the production rate by the sustainable practices enabled by the processing and the analysis of industrial data using Iot and big-data.

References

[1] De Mauro, Andrea, Marco Greco, and Michele Grimaldi. "What is big data? A consensual definition and a review of key research topics." In *AIP conference proceedings*, vol. 1644, no. 1, pp. 97-104. AIP, 2015.

[2] Jones, Matthew. "What we talk about when we talk about (big) data." *The Journal of Strategic Information Systems* 28, no. 1 (2019): 3-16.

[3] Mikalef, Patrick, Ilias O. Pappas, John Krogstie, and Michail Giannakos. "Big data analytics capabilities: a systematic literature review and research agenda." *Information Systems and e-Business Management* 16, no. 3 (2018): 547-578.

[4] Tonidandel, Scott, Eden B. King, and Jose M. Cortina. "Big data methods: Leveraging modern data analytic techniques to build organizational science." *Organizational Research Methods* 21, no. 3 (2018): 525-547.

[5] Sun, Zhaohao, Lizhe Sun, and Kenneth Strang. "Big data analytics services for enhancing business intelligence." *Journal of Computer Information Systems* 58, no. 2 (2018): 162-169.

I-SMAC

[6] Amado, Alexandra, Paulo Cortez, Paulo Rita, and Sérgio Moro. "Research trends on Big Data in Marketing: A text mining and topic modeling based literature analysis." *European Research on Management and Business Economics* 24, no. 1 (2018): 1-7.

[7] Storey, Veda C., and Il-Yeol Song. "Big data technologies and management: What conceptual modeling can do." *Data & Knowledge Engineering* 108 (2017): 50-67.

[8] Akter, Shahriar, Samuel Fosso Wamba, Angappa Gunasekaran, Rameshwar Dubey, and Stephen J. Childe. "How to improve firm performance using big data analytics capability and business strategy alignment?" *International Journal of Production Economics* 182 (2016): 113-131.

[9] Akter, Shahriar, and Samuel Fosso Wamba. "Big data analytics in E-commerce: a systematic review and agenda for future research." *Electronic Markets* 26, no. 2 (2016): 173-194.

[10] Sivarajah, Uthayasankar, Muhammad Mustafa Kamal, Zahir Irani, and Vishanth Weerakkody. "Critical analysis of Big Data challenges and analytical methods." *Journal of Business Research* 70 (2017): 263-286.

[11] Hu, Jiankun, and Athanasios V. Vasilakos. "Energy big data analytics and security: challenges and opportunities." *IEEE Transactions on Smart Grid* 7, no. 5 (2016): 2423-2436.

[12] Lv, Zhihan, Houbing Song, Pablo Basanta-Val, Anthony Steed, and Minho Jo. "Next-generation big data analytics: State of the art, challenges, and future research topics." *IEEE Transactions on Industrial Informatics* 13, no. 4 (2017): 1891-1899.

[13] Zhou, Keliang, Taigang Liu, and Lifeng Zhou. "Industry 4.0: Towards future industrial opportunities and challenges." In *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pp. 2147-2152. IEEE, 2015.

[14] Kiel, Daniel, Julian M. Müller, Christian Arnold, and Kai-Ingo Voigt. "Sustainable industrial value creation: Benefits and challenges of industry 4.0." *International Journal of Innovation Management* 21, no. 08 (2017): 1740015.

[15] Sun, Yunchuan, Houbing Song, Antonio J. Jara, and Rongfang Bie. "Internet of things and big data analytics for smart and connected communities." *IEEE access* 4 (2016): 766-773.

[16] Witkowski, Krzysztof. "Internet of things, big data, industry 4.0–innovative solutions in logistics and supply chains management." *Procedia Engineering* 182 (2017): 763-769.

[17] Koo, Dan, Kalyan Piratla, and C. John Matthews. "Towards sustainable water supply: schematic development of big data collection using internet of things (IoT)." *Procedia engineering* 118 (2015): 489-497.

[18] Maksimovic, M. "Greening the future: Green Internet of Things (G-IoT) as a key technological enabler of sustainable development." In *Internet of things and big data analytics toward next-generation intelligence*, pp. 283-313. Springer, Cham, 2018.

[19] Channe, Hemlata, Sukhesh Kothari, and Dipali Kadam. "Multidisciplinary model for smart agriculture using internet-of-things (IoT), sensors, cloud-computing, mobile-computing & big-data analysis." *Int. J. Computer Technology & Applications* 6, no. 3 (2015): 374-382.

I-SMAC

[20] Mourtzis, D., E. Vlachou, and N. Milas. "Industrial Big Data as a result of IoT adoption in manufacturing." *Procedia Cirp* 55 (2016): 290-295.

I-SMAC