JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

COMPREHENSIVE REVIEW OF USING BIG DATA ANALYTICS FOR SMART SEMICONDUCTOR MANUFACTURING

[1]Arpit Tripathi, [1]Soniya Singh, Srivaramangai. R[2]
[1]Department of Information Technology, University of Mumbai
[2]Department of Information Technology, University of Mumbai

Abstract : The emergence of Industry 4.0 is to optimize production processes, increase efficiency and reduce costs[1]. The integration of big data analytics into smart design enables the collection and analysis of large volumes of data from multiple sources, including sensors, machines, and human input. This article provides a qualitative review of the existing literature on big data analytics in smart manufacturing, including challenges, implications, and future directions. We examine different methods and techniques used in big data analysis, such as machine learning, deep learning and natural language processing, and their applications at different levels of the manufacturing process, from product to after-sales service. We also discuss issues related to the use of big data analytics in smart design, including data quality, data security, and data management. Finally, we present future research directions and recommendations for the use of big data analytics in intelligent design, such as the development of hybrid algorithms and the integration of time analysis. This study will be beneficial for researchers, practitioners and policy makers who want to use big data analytics for smart manufacturing.

***IndexTerm* – Smart Manufacturing, Semiconductor, Big Data, Data Analysis, Big Data Architecture**

I. INTRODUCTION

Smart Manufacturing is a term often used for the movement in the production of upstream and downstream integration in the supply chain, cyber integration of physical and capacity, and use of advanced information to be flexible and resilient [1]. It is often equated with "Industry 4.0", a term coined by the German government's attempt to promote fourth-generation production using techniques such as cyber-physical systems, real and process copies of equipment, and decision making. To create smart factories SM uses the same production process in the volume, speed, and diversity of data (often referred to as "big data") and uses "big data analytics" to improve existing resources and provide new resources such as predictive analytics. These resources are common across all industries, but in varying proportions, perhaps due to factors such as infrastructure, culture, supply base, and demand. Therefore, there is a potential for cross-industry analysis that can be used to improve resources in certain industries such as biochemistry and biology. This article explores recent developments in SM big data targeting the semiconductor industry to identify potential applications in biochemistry. In particular, the following sections describe the history of semiconductor industry factory operations, the emergence of SM, and plans for the use of big data analytics. Analytical methods in semiconductor manufacturing then issues of managing the development of manufacturing analytics, understanding talent analytics taxonomy, current state. This art test and analytical roadmap support the SM concept. Application case studies are then presented, highlighting the capabilities of analytical methods and the SM models they support. The article ends with a discussion of the key concepts guiding the development of big data analytics for semiconductor fabrication; suggests that these concepts will have implications in biochemistry and other industries. The appendices provided include a list of abbreviations used in the document and instructions for various analytical methods and solutions.

1.1. Smart Manufacturing Analysis [2]

1.1.1. Understanding Business

This Section focuses on understanding smart manufacturing concepts in consultation with experts. The purpose of the project is derived from Dolle's view of what needs to be done, which is then translated into a case study. From a business perspective, some of Dolle's main goals are described below. Once the machine worked, Dolle wanted to know how long it would take to reach a reasonable price for the product. Business objectives specify objectives in the business context, while data mining objectives specify objectives in context. A non-exhaustive list of data mining purposes is below. What is the frequency and total downtime of machine stops due to incorrect rope/screw error? How fast is the object moving through the machine? What is the maximum speed? Is work in pace?. Estimate machine uptime and how it will be maintained, based on historical standards? What is the total time of the machine? What's the price? In general, production is considered efficient with 80-85% efficiency. It makes sense to investigate all expected and unknown problems/challenges in the production process [3].

1.1.2. Understanding Big Data

Big data must first be understood in order to be used in a smart manufacturing project. Big data describes vast amounts of information that are produced quickly and in a variety of ways and formats. These data sources include sensors, machines, social media, and other digital channels. They might be structured or unstructured. Big data analytics uses sophisticated analytical methods and tools to glean insightful information from this enormous body of data.

Big data analytics can be employed in a smart manufacturing project in several ways, including:

- **Predictive maintenance:** Manufacturers can identify when a machine is most likely to break down and take preventative action to avoid downtime and lower maintenance costs by gathering and analysing real-time data from equipment and sensors.

- **Quality control:** Manufacturers can spot quality problems early on and take corrective measures to enhance product quality by analysing production data.

- **Supply chain optimization:** Manufacturing companies can decrease costs and increase efficiency by optimising their supply chains by analysing data from their suppliers.

- **Energy optimisation:** Manufacturers can find chances to cut energy use and optimise their energy consumption by analysing data on energy use.

Manufacturers must make the necessary infrastructure investments, including those in data processing, storage, and analytics tools, in order to integrate big data analytics in a smart manufacturing project. Additionally, they must make sure the data is accurate and pertinent to their manufacturing processes. They must also make sure that their analytics tools are scalable and able to keep up with the expansion of data over time.

1.1.3. Data Preparation

This section provides an understanding of business considerations before processing data models. The data preparation phase includes activities such as data selection, data transfer, data cleaning and data validation. Data preparation tasks can be done multiple times and not in any order. Currently, important issues such as selecting and cleaning relevant data, discarding invalid data and how to integrate the ERP data system into the latest information are discussed. Through discussions between data scientists and domain experts, metadata history demonstrates importance in the data validation process.

1.1.4. Modelling

This chapter introduces the main concepts of model-based machine learning and explains some important topics such as modelling. One of the main goals of this research is to predict the unplanned nature of machines based on historical events/patterns. Depending on the type of data available and the question/purpose, machine learning can be used to predict when the machine will stop. Tracking learning algorithms are trained on historical data, as in the "1" or "0" state of the machine. The algorithm determines which new label should be assigned based on the past model.

1.1.5. Analysis

This section evaluates the OEE of the Dolle manufacturing process. OEE calculates the percentage of production hours produced. It can be used as a benchmark and as a basis. In general, OEE has three aspects: usability, performance, and quality. The account has all the events that are no longer scheduled. Efficiency takes into account the factors that cause the production process to run at less-than-ideal speed. According to the agreement, products that do not meet the quality standards are evaluated. An OEE score of 100% means that production is running optimally without any unplanned downtime and only quality products are produced.

1.2. Semiconductor Manufacturing Background

1.2.1. Semiconductor Manufacturing and Emergence of Online Analytics

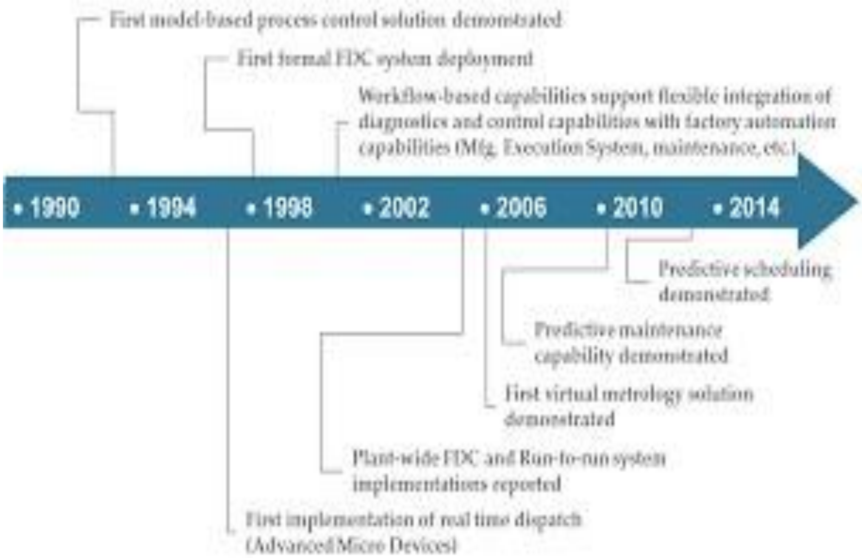

figure 1: semiconductor manufacturing and emergence of online analytics

1.2.2. The Big Data Revolution and Related Challenges

Roadmap to Emerging Big Data in Semiconductor Operations is defined in "5 Vs" Volume, Velocity, Variety (or data collection), Veracity (or data quality), and value (or clinical utility). In terms of various information, information storage such as equipment tracking information for measurement, maintenance, distribution, inventory management, production process (MES) and enterprise resource planning (ERP) has been used for over a decade; however, the review is now active. supporting and integrating multiple stores to discover relationships, identify anomalies and predict events. Data quality (veracity) is considered the most important "V" data, reflecting the widespread use of big data analytics in business. Accuracy, completeness, content richness, accessibility, and archive length were identified as areas where the industry needs to improve data quality to support this advanced analysis. The development of big data in terms of value has led to the development of existing analytical applications and the emergence of new applications.

1.2.3. Drawing from other industries

Semiconductor manufacturing has resource development potential in other industries such as biochemistry and biology.

II. Literature review

Intelligent manufacturing is a term often used to increase productivity through the use of data, including collaboration, the combination of physical and cyber capabilities, and changes in the use of big data. SM adoption is uneven across the industry, so there is an opportunity to look to other industries to identify solutions and strategic plans for industries such as biochemistry or biology.[1] The development of big data provides the opportunity to manage large volumes of data and perform through analytics to improve diagnosis and prognosis. The analysis process can be interpreted in terms of dimensions to understand their needs and capabilities and identify gaps in technology[2]. The semiconductor industry has benefited from advances in big data and analytics by improving existing functions such as fault detection and enabling new functions such as predictive maintenance. For most of these projects:

- Good data is the most important data for delivering good solutions.
- It is necessary to include skills in the analysis to arrive at a good online solution[3].

In the future, analysis will be further supported by the development of the big data environment with the creation of good ideas such as digital twins; however, he thinks that the integration of these skills will still be needed in the design process[4].

2.1. Big Data Broadcasting in Semiconductor Manufacturing

Next Generation Fault Detection and Classification (NG-FDC)

Failure detection has been an integral part of semiconductor manufacturing for at least one year Tracked and has significant benefits (such as reducing the amount of spare parts), improved quality and maintenance guidance), is also affected by the high installation and high cost of bad and bad. Here, the top APC experts in the microelectronics design community of customers and vendors agree on the following points:

- "From design to the automation of the FD front-end for constraint management [has], but management process and intellectual property"; and
- A major problem in today's FDC systems is negative and/or false positives. It has been spotted in a real place.

The development of big data offers semiconductor manufacturing an opportunity to provide advanced FDC capabilities to address these critical issues. The FD part of NGFDC technology has two main points: The first prevention is row-level analysis: this equipment uses data-driven multivariate analysis techniques to identify and show inconsistencies. It has the advantage of an easy-to-manage analysis method. It is also possible to identify patterns in the data that are not obvious or acceptable for the SME. However, without prior warning, etc. can cause many negative and negative effects. The second part of, Semi-Automatic Part Segmentation, Feature Extraction, and Limit Tracking, directly addresses the trouble spots identified at the APC Council business meeting cited above.

The method uses semi-automatic trajectory segmentation, feature extraction and constraint detection techniques that allow SMEs integration. Using this method, the design time can be reduced several times depending on the SME's level of integration, and the results will improve both negatively and negatively. The distribution line is shown in Figure 2. The relevant model was determined by analysing historical data and consulting SMEs. In fact, many of these patterns are usually signals produced by physical processes such as open switches (cascade continuous operation), vibrating or low-damping drivers (oscillations), interference (spikes) or drifts (ramps) for periods.

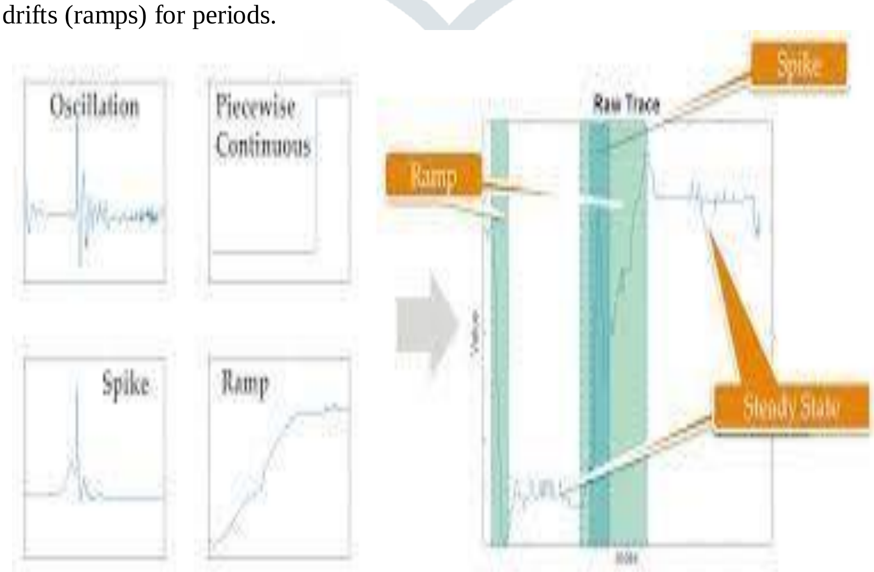

Figure 2 The relevant model

Three analysis methods were compared:

1. Data analysis as a whole using all tools and variances to analyse errors.
2. The Grid window is best used for estimation grid segmentation and feature extraction; and
3. The above is used for semi-automatic feature extraction FD. Group

SOM was used to develop alarm limit settings for each route.

The results show that the proposed semi-automatic trajectory segmentation and feature extraction outperforms other methods in reducing false positives and negatives.

2.2. Big Data Transformation and Analytics

2.2.1. Bigdata Architectures [6]

The move to Big Data analytics to support the SM strategy in semiconductor manufacturing includes the move to more Big Data friendly systems. For example, Apache Hadoop is an open-source software framework for storing and processing data at scale on a set of hardware devices. Hadoop leverages parallel processing and scalability capabilities to provide tailored solutions for large time series datasets such as data monitoring. Most of the big semiconductor companies are integrating or at least measuring big data, the main choice is Hadoop.

Overall, migrating to a Hadoop infrastructure offers many benefits, including:

• Lower cost of ownership (COO) for data solutions.
• More data storage.
• Improved query speed.
• Improve performance appraisal; and
• Optimizing predictive analytics that measures these results

However, these platforms are generally not suitable for real-time or real-time online analytics on the factory floor. Therefore, the future factory data management solution will likely be Hadoop, or another big data friendly ecosystem combined with small business products that will be best suited for data. Important moments (e.g., "real time") analysis and application [7].

2.2.2. Emergence of artificial intelligence and other big data friendly analytics

The term "AI" can be used to describe any device or analysis that is aware of its environment and functions to achieve its purpose. Today, the word is often used to refer to this concept of device or analysis that monitors the work of the human brain, such as in self-driving cars. Artificial neural networks (ANNs) are an example of such analysis and have been used for decades. This type of AI analytics is remerging as part of the big data revolution. For example, deep learning is a technique similar to traditional ANNs that uses a level of abstraction to improve the quality and speed of analysis of large data sets. “Deep learning can be used to solve some of the most important problems in big data analysis, such as models generated from big data, measurement analysis, data logging, fast data storage, and easy separation. Also, designs are often impractical, so it is often difficult to assess the future of these designs. There is a lot of research soon focused on engaging SMBs with AI strategies; these strategies should be used in the future [8]. The use of big data is content analysis with solutions often referred to as "navigators". These applications do data mining in the background, looking for patterns or statistics of interest, such as the location of nearby faults, and then asynchronously notifying the user of what needs to be done, such as the management of facilities. This approach allows for background scanning so that they can be better controlled by more important tasks such as weed control. It also allows the development and management of diagnosis and prognosis for rehabilitation[9].

2.2.3. Achieving the smart manufacturing vision

- Integration with the supply chain: Smart manufacturing will become an integral part of the supply chain and the optimization of the Factory will be a part of the entire supply chain optimization.
- Enhanced use of Cyber-Physical Systems (CPS): CPS stands for "the relationship and coordination between computing and physical systems".
- Combining the simulation environment with the context of virtual factories and digital twins: Predicting the vision for semiconductor manufacturing, "(1) Yield and Throughput Forecasting is an important part of Fab Operations and
- (2) Real-time simulation of all factory operations, existing systems it dynamically updates the simulation model.”
- Infrastructure Architecture with Big Data: This infrastructure includes analysis and applications that use methods to improve SM performance as well as data management.
- Leverage advanced analytics: As described in this document, the benefits of Big Data evolution will primarily be achieved through the use of analytics that use big data to improve existing resources (like FDC) or develop new capabilities (like forecasting))[10].

III. Results and Future Work

This article presents the main concepts of data analysis based on real research data. These topics include data understanding, data preparation, data pipelines, and data analysis techniques. In-depth analysis and forecasting are done to improve performance. Monitoring machine learning is used to develop classification models to predict machine braking. The results highlight the improvement of production by reducing machine time. The predictions made by the model are perfectly acceptable in estimating the decline, which is one of the main causes of the deterioration in production. Intelligent design is the evolution of design capabilities that focuses on improving performance through tighter integration, physical connectivity and cyber capabilities, leveraging the evolution of big data, and leveraging data and analytics. This change happens unevenly across the industry. Therefore, there is an opportunity to look to other industries to help identify solutions and methods for specific industries such as biochemistry or biology. Analysis and comparison of these analyses using a qualitative method allows to identify effectiveness as well as gaps and opportunities for future improvement. Semiconductor manufacturing is characterized by the enforcement of regulations, the equipment and procedures are very complex, and the quality of materials. Analytics will continue to be used at a higher rate in VM and will generally correspond to the evolution of Big Data. These factors will continue to be important in the SM analytical roadmap, although methods for developing good knowledge and SME integration will evolve as issues are resolved. For future studies, various predictive machine learning models will be used and compared. In addition, a near real-time dashboard will be created to display I/O speed and OEE data. Finally, it will examine how descriptive analytics, predictive analytics, and near real-time dashboards can help the company's intelligent design environment improve operations and productivity.

IV. REFERENCES

[1] Xu, L. D., Xu, E. L., & Li, L. (2018). Industry 4.0: state of the art and future trends. International Journal of Production Research, 56(8), 2941-2962.

[2] Yen, J., Chen, W., & Wu, J. (2018). Big data analytics for smart manufacturing. Journal of Industrial and Production Engineering, 35(6), 357-360.

[3] Zhang, M., Zhou, M., & Hu, J. (2020). A big data analytics framework for smart manufacturing based on internet of things. Journal of Intelligent Manufacturing, 31(1), 93-106.

[4] Zhu, L., Deng, Z., Li, S., Li, B., & Li, Y. (2019). A review of data-driven approaches for smart manufacturing. Journal of Manufacturing Systems, 50, 38-48.

[5] May, G.S.; Spanos, C.J. Fundamentals of Semiconductor Manufacturing and Process Control; IEEE, Wiley-Interscience: Hoboken, NJ, USA, 2006.

[6] Suschitzky, E., Markus, H., and Agesan, R. (2014). How big data can improve manufacturing. Available online at: https://www.mckinsey.com/businessfunctions/operations/our’s-insights/ how-big-data-can-improve-manufacturing.

[7] Romero-Torres, S.; Moyne, J.; Kidambi, M. towards Pharma 4.0; Leveraging Lessons and Innovation from Silicon Valley. American Pharmaceutical Review, 5 February 2017.

[8] Rebooting the IT Revolution: A Call to Action; Semiconductor Industry Association and Semiconductor Research Corporation: Washington, DC, USA, 4 January 2016.

[9] Chen, H., Mao, S., & Liu, Y. (2018). Big data analytics for smart manufacturing: Case studies in semiconductor manufacturing. Journal of Advanced Manufacturing Systems, 17(02), 181-192.

[10] Jin, H., Zhang, C., & Yu, X. (2019). Big data analytics for smart manufacturing: a review. Journal of Industrial Information Integration, 13, 1-11.

[11] Wang, S., Wan, J., Zhang, D., Li, D., & Zhang, C. (2016). Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. Computer Networks, 101, 158-168.

[12] Wei, J., Wu, X., Zhao, L., & Yang, Y. (2020). Real-time monitoring and big data analytics for intelligent machining of aero-engine blades. Robotics and Computer-Integrated Manufacturing, 61, 101839.

[13] Dey, P., Sharma, V., & Ashour, A. S. (2019). Big data analytics: A review on theoretical contributions and applications in social network, business and healthcare. Journal of Business Research, 98, 389-403.

[14] Li, Q., Li, X., Liang, X., & Huang, X. (2018). Big data analytics in healthcare: a literature review. BioMed Research International, 2018, 1-13.

[15] Oussous, A., Benjelloun, F. Z., Lahcen, A. A., & Belfkih, S. (2018). Big data analytics: A survey. Journal of King Saud University-Computer and Information Sciences, 30(4), 431-448.

[16] Palanisamy, V., Gerasimov, V., & Buyya, R. (2018). Big data analytics and its ecosystem: A systematic literature review of research trends. IEEE Access, 6, 18084-18106.

[17] Sharma, R., & Bhatia, V. (2019). Big data analytics for financial services: A review. Journal of Business Research, 103, 219-230.

[18] Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data mining with big data. IEEE Transactions on Knowledge and Data Engineering, 26(1), 97-107.