IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 4, April 2024

Enhancing Quality in Industry 4.0: A Data-Centric Approach to Lean Six Sigma

Saarthak Chawla
Student, Department of Computer Science,
Dronacharya College of Engineering, Gurgaon, India

Abstract*: Nearly all methods aimed at improving quality rely on the collection and analysis of data to address quality issues. The fusion of six sigma and lean manufacturing results in lean six sigma methodology, which targets achieving six sigma quality levels (less than 3.4 defects per million) by minimizing variations and inefficiencies in processes. Attaining this objective hinge on meticulous data collection to tackle quality challenges.*

While many conventional data analysis techniques are applicable for enhancing product and process quality, the advent of Industry 4.0 technologies generates vast datasets that necessitate robust data analysis methods to derive actionable insights from big data. Employing these analysis methods throughout the lean six sigma cycles, particularly during the measurement and analysis phases, is crucial for making informed decisions.

This study aims to offer a comprehensive guide for implementing lean six sigma to expedite decision-making processes, enhance reliability, and foster satisfaction through data utilization. It not only enhances manufacturing processes by reducing lead times and delivering superior quality products but also facilitates effective decision-making through various mining techniques..

Keywords: Lean Six Sigma, Data Analysis, Efficiency, Quality

I. INTRODUCTION

Improving product quality is paramount in industries where numerous competitors offer similar products. To enhance quality, gathering extensive data and extracting valuable insights from it are crucial steps. Various quality improvement methods, including inspection, statistical process control, total quality control, zero defects, kaizen, and lean six sigma (LSS), rely on data collection to address quality issues (Köksal, et al., 2011).

LSS emerged as a fusion of two distinct methodologies initially, namely six sigma and lean manufacturing. Despite their differences, these methodologies possess complementary features that make their integration highly effective for achieving sustainable operational outcomes. Six sigma is renowned for its systematic problem-solving approach and methodology, widely adopted across diverse industries (Lee, et al., 2004; General Electric, 2017). It utilizes statistical measures, with "sigma" representing the degree of deviation of a process from perfection, aiming to reduce errors in manufacturing to an exceptionally low level of 3.4 parts per million (ppm), ultimately striving for zero defects.

Conversely, lean manufacturing focuses on maximizing value-creating operations while eliminating non-value-added activities from the customer's perspective. By identifying and eliminating waste, lean manufacturing reduces costs and shortens lead times. Although lean manufacturing and six sigma differ in their approaches to identifying the root causes of waste, both aim to produce high-quality products that meet customer expectations effectively. The overarching goal of both methodologies is to create an efficient production system that minimizes defects and waste.

LSS has garnered widespread popularity due to its ability to deliver both output and quality-based improvements. By systematically simplifying manufacturing processes and striving for less than 3.4 ppm errors, LSS has led to significant enhancements and cost savings in various companies such as General Electric (2017), Dell Inc. (Lovin and Yaptangco, 2006), and Xerox Corp. (2004).

About 95% of LSS initiatives focus on enhancing quality through the Define-Measure-Analyze-Improve-Control (DMAIC) approach (Trnka, 2012; Kanakana, et al., 2010). This methodology aims to achieve six sigma quality levels, with fewer than 3.4 defects per million opportunities. DMAIC, as demonstrated by the Six Sigma system, provides a structured framework for problem-solving, identifying future opportunities, and effectively managing projects. Widely regarded as one of the most efficient problem-solving methods to date, DMAIC mandates the use of data for various purposes (George, et al., 2004).

While most Six Sigma projects utilize the DMAIC cycle, some opt for the DMADV cycle, also known as Design for Six Sigma, particularly for creating new processes. While the Define, Measure, and Analyze phases of DMADV resemble DMAIC, they serve different purposes. For instance, Huang et al. (2010) applied DMADV to enhance product quality, while Chen et al. (2005) utilized it to improve the assembly efficiency of military products. Table 1 provides a summary of the steps involved in both DMAIC and DMADV cycles, along with sample activities corresponding to each step.

TABLE I: Six Sigma Process Improvement Framework: Phases, Definitions, and Procedure

Phases	**Definitions**	**Steps to be followed**
Define	Define the project's purpose, scope, and goals.	Identify customer requirements and project objectives. Develop a project charter. Formulate a high-level process map.
Measure	Measure current process performance and gather data.	Define critical-to-quality (CTQ) parameters. Collect relevant data. Analyze data to understand current performance.
Analyze	Analyze data to identify root causes of issues.	Use statistical tools to analyze data. Identify root causes of defects or variations. Prioritize potential causes for further investigation.
Improve	Improve the process by implementing solutions.	Generate and test potential solutions. Implement changes to the process. Monitor the effects of changes and adjust as necessary.
Control	Sustain improvements and ensure process stability.	Develop control plans to maintain improvements. Implement monitoring systems. Establish procedures for addressing deviations and continuous improvement.

Achieving accurate results with Lean Six Sigma (LSS) entails a systematic reduction of waste and costs (Polk, 2011), emphasizing cycles based on data. With the advent of Industry 4.0 technologies, there is an abundance of data available for collection. Utilizing various mining techniques like big data analytics, data mining, and process mining becomes crucial. These techniques enable decision-makers to uncover insights not immediately apparent, saving time and ensuring decisions are data-driven. Despite widespread adoption of lean manufacturing and Six Sigma, many companies remain dissatisfied with outcomes (Guarraia et al., 2008). Modelling and optimizing processes pose challenges due to the volume of quality-related data. In response, decision-makers are compelled to store extensive data to avoid erroneous decisions. However, in the Industry 4.0 era, this abundance of data increases the risk of flawed decisions even as it becomes indispensable for informed decision-making. Leveraging big data techniques like text mining, machine learning, and deep learning, alongside fundamental data mining techniques such as clustering and prediction algorithms, assists in making correct and optimal decisions across various stages of LSS. Thus, integrating mining techniques into LSS cycles is essential for effective decision-making in businesses.

II. IMPORTANCE OF DATA MINING TECHNIQUES IN LEAN SIX SIGMA (LSS)

In LSS, data plays a crucial role, and mining this data for meaningful insights becomes paramount. As technology evolves, new methods of mining data emerge, driven by the vast amounts of data collected. Traditional data mining techniques prove effective for smaller datasets, but with the advent of automatic data collection through sensors, big data analytics becomes preferable due to the sheer volume of data available. Additionally, process mining methods gain significance in LSS as they focus on enhancing process quality. It is essential to remember that the purpose of analytics

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 4, April 2024

is to extract insights from data, rather than evaluating the entire process. Analytics can aid in various LSS activities, such as identifying root causes and developing robust solutions, but it is only a part of the LSS process.

Data Preparation and Evaluation

Data preparation and evaluation are essential steps in all mining techniques. Data preprocessing is necessary due to the redundancy, incompleteness, and inconsistency of collected data. Tasks such as data cleaning, transformation, reduction, and discretization are vital to enhancing data quality before mining. Evaluating mining results involves comparing different techniques, selecting the one that yields the best outcome, and ensuring that the knowledge extracted from datasets is correctly interpreted for making optimal decisions.

Data Mining Techniques

Data mining is an exploratory data-analytic process that uncovers interesting patterns within large datasets. These techniques utilize integrated data stored in databases using statistical and mathematical methods. They can be categorized into descriptive and predictive tasks, including association analysis, clustering, classification, and prediction, all of which play significant roles in LSS cycles.

Big Data Analytics

Big data analytics deals with large volumes of varied data generated, captured, and processed at high velocity. With the exponential growth of data from various sources, accurate decision-making requires real-time analysis. Big data analytics, including machine learning, text mining, and video mining, provide deep insights into processes, complementing LSS efforts. These techniques help in analysing data from sources like sensors, smart devices, and organizational systems to make informed decisions.

Process Mining

Process mining focuses on extracting process-centred insights by analysing event logs from information systems. It aids in discovering, tracking, and improving real transactions to eliminate variations and wastes in processes. Process discovery, conformance checking, and enhancement are key aspects of process mining, contributing to the optimization of LSS efforts. Unlike traditional data mining, process mining is process-centric and addresses questions related to performance and compliance, making it a valuable tool in LSS.

III. A PROPOSED FRAMEWORK FOR QUALITY IMPROVEMENT

In Figure 1, a flow diagram illustrates the stages of Lean Six Sigma (LSS) for both existing (DMAIC) and new (DMADV) processes.

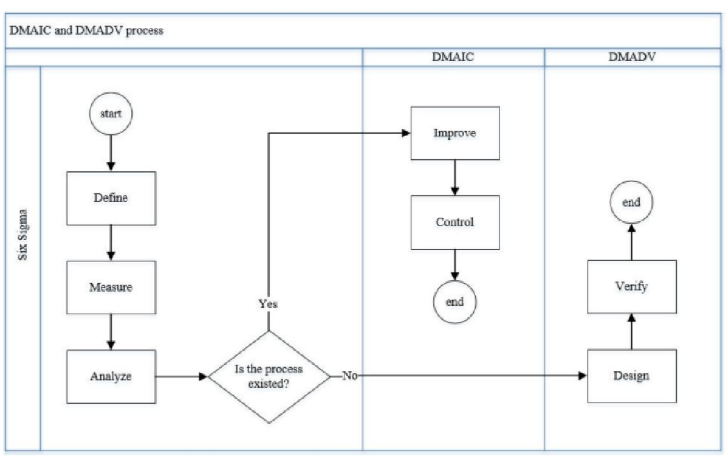

Fig. 1. Stages for Lean Six Sigma Processes

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 4, April 2024

The Define phase is nearly identical in both cycles. While there are some differences in the Measure and Analyze phases, they can be considered similar overall, with no significant distinctions. In existing processes, the Improve phase is crucial for making informed decisions to enhance quality, followed by the implementation of control measures. Conversely, when creating a new process, the focus shifts to the design phases, ensuring that the proposed process is thoroughly developed and validated before implementation.

TABLE III: Methods Used in Lean Six Sigma (LSS) Cycle

Phases	Method	Description
Define	Brainstorming	Group technique for generating creative ideas or solutions to potential problems.
	Nominal Group Technique	Structured method for group decision-making, ensuring equal participation and minimizing biases.
	Prioritization Matrix	Tool for ranking and prioritizing issues or problems based on criteria such as impact and feasibility.
	Pareto Analysis	Technique for identifying and prioritizing the most significant factors contributing to a problem, typically using the 80/20 rule
	Quality Function Deployment (QFD)	Methodology for translating customer requirements into specific product or process design characteristics.
	Big Data Analytics	Utilization of advanced analytics techniques to analyze large volumes of data, including unstructured data such as text and video, for insights and decision-making.
Measure	Data Collection	Systematic gathering of relevant data points to assess current process performance and identify areas for improvement.
	Statistical Process Control (SPC)	Application of statistical methods to monitor and control a process, ensuring it operates within specified limits and meets quality standards.
Analyse	Root Cause Analysis	Systematic approach for identifying the underlying causes of problems or defects in a process.
	Regression Analysis	Statistical method for determining the relationship between one or more independent variables and a dependent variable.
Improve	Design of Experiments (DOE)	Methodology for systematically varying factors to identify the optimal combination for achieving desired outcomes.
	Failure Mode and Effects Analysis (FMEA)	Proactive technique for identifying and mitigating potential failures or defects in a process or product.
Control	Control Charts	Graphical tools for monitoring process performance over time and detecting any deviations from the desired standard.
	Standard Operating Procedures (SOPs)	Documented instructions detailing the steps and responsibilities required to carry out specific tasks or processes consistently.

In the DMADV cycle, the Define phase involves identifying problems and establishing process goals aligned with customer needs. Table 2 presents various methods applicable to both DMAIC and Define phases, aiding in this initial stage. For instance, brainstorming and the nominal group technique are valuable for defining potential problems. Once potential problems are identified, methods such as the prioritization matrix or Pareto analysis help determine which issue(s) to address first. Quality Function Deployment (QFD) proves beneficial in establishing process goals based on customer requirements. Additionally, for unstructured data like text and video, big data analytics can assist in problem definition. Process discovery offers a visual representation of the process status, aiding in the Define phase. Figure 2 outlines the recommended roadmap for the Define phase.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 4, April 2024

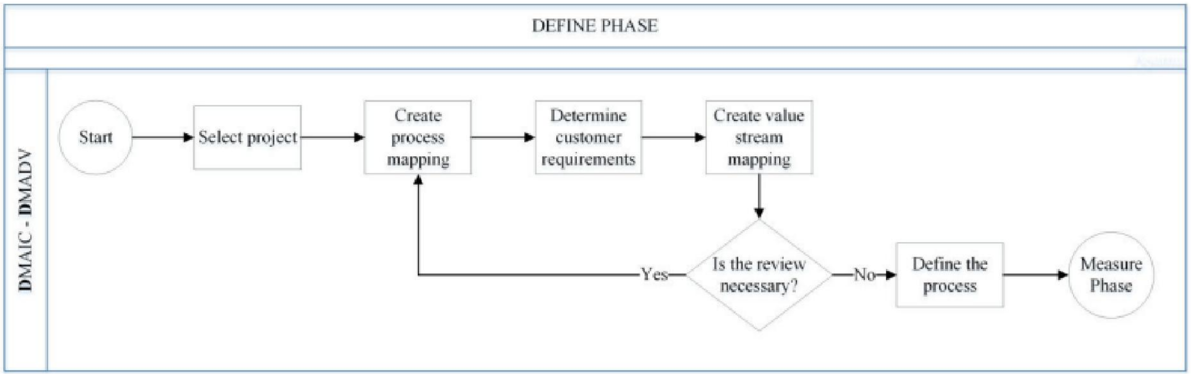

Fig. 2. Lean Six Sigma - Define

One of the primary objectives of Lean Six Sigma (LSS) is to mitigate variations within the process. During the Measure phase, analysis of variance (ANOVA) serves to pinpoint the sources of variation leading to quality issues. Additionally, tally charts, along with descriptive and predictive statistics, offer valuable insights into the current state of the process. Figure 3 provides a visual representation of the recommended steps for the Measure phase.

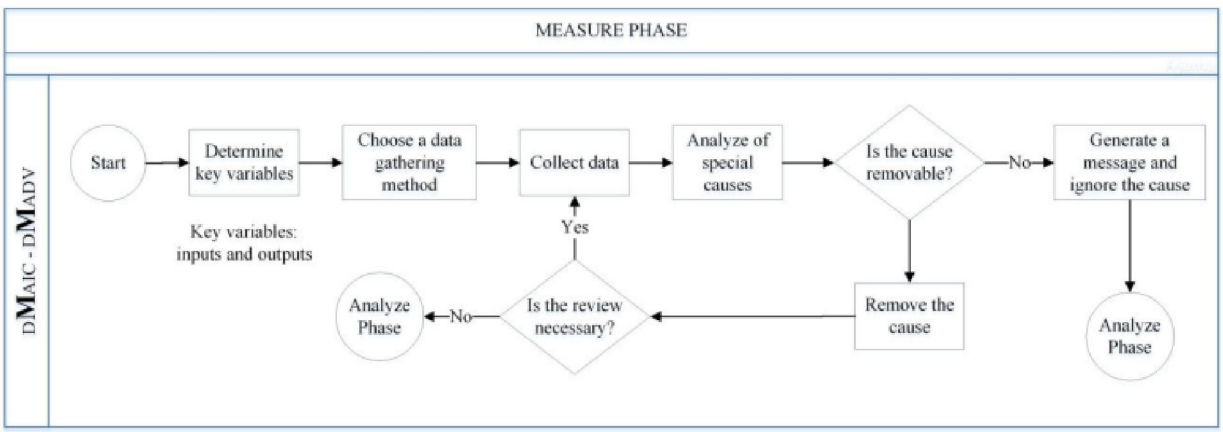

Fig. 3. Lean Six Sigma - Measure

Data collected in the Measure phase is then utilized for process analysis. Correlation and regression analyses are employed to elucidate the relationship between dependent and independent variables. Process capability and performance are evaluated using Statistical Process Control (SPC). Among data mining (DM) techniques, association rules, clustering, and classification are commonly employed to analyze processes. Machine learning algorithms facilitate user navigation without the need for extensive learning efforts. Moreover, if the dataset includes timestamps, process discovery, a type of process mining (PM), proves to be an efficient technique for process analysis. Figure 4 outlines the suggested roadmap for the Analyze phase.

Data collected during the Measure phase serves as the basis for process analysis. Correlation and regression analyses are employed to elucidate the relationships between dependent and independent variables. Evaluation of process competence and performance is conducted using Statistical Process Control (SPC). Association rules, clustering, and classification are among the most favored data mining (DM) methods for process analysis. Utilizing machine learning algorithms can facilitate user navigation without requiring extensive learning efforts. Moreover, if the dataset includes timestamps, process discovery, a form of process mining (PM), may offer a more efficient technique for process analysis. Figure 4 illustrates the suggested roadmap for the Analyze phase.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 4, April 2024

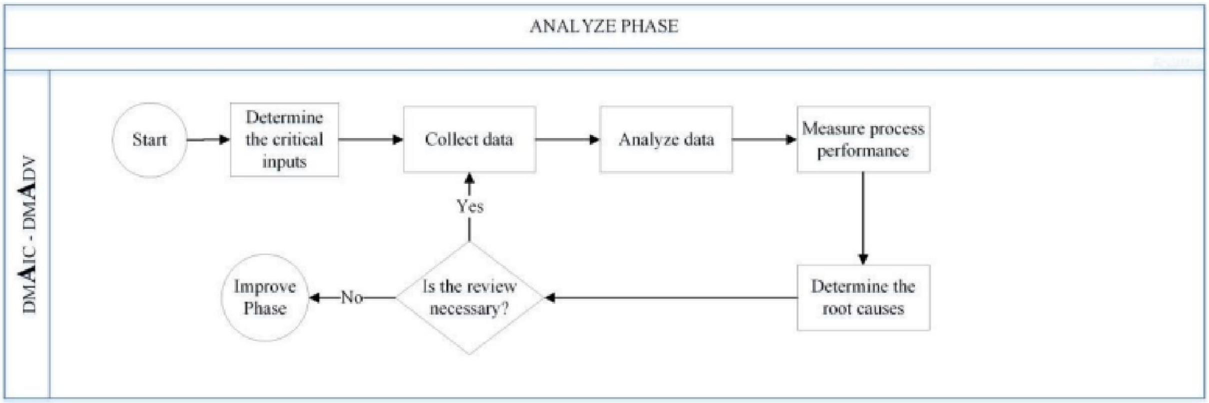

Fig. 4. Lean Six Sigma - Analyze

The Improve phase can benefit from the application of multivariate statistical methods such as clustering analysis and discriminant analysis to enhance the process. Techniques like TRIZ can provide innovative solutions to complex problems, while Design of Experiments (DOE) aids in identifying the optimal parameters for process improvement. Figure 5 presents the suggested roadmap for the Improve phase.

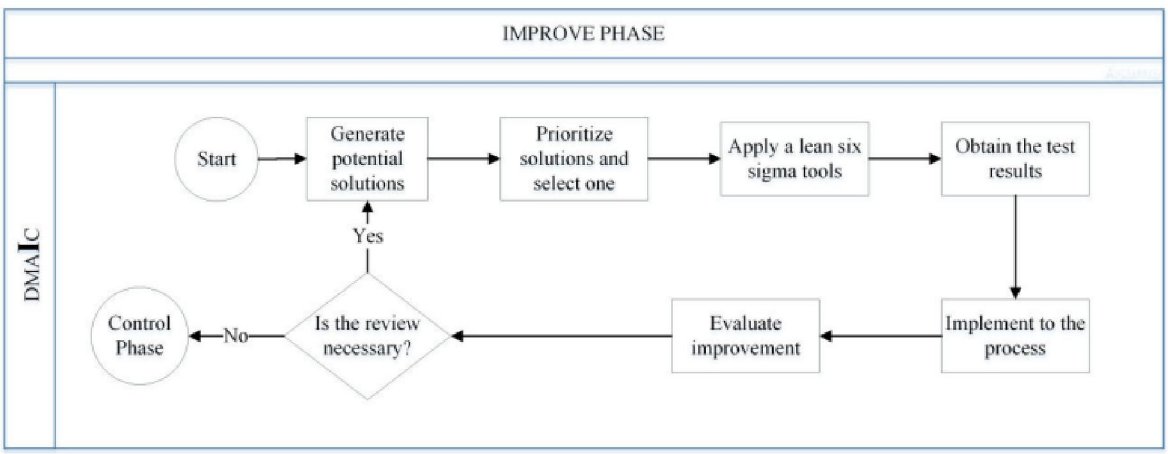

Fig. 5. Lean Six Sigma - Improve

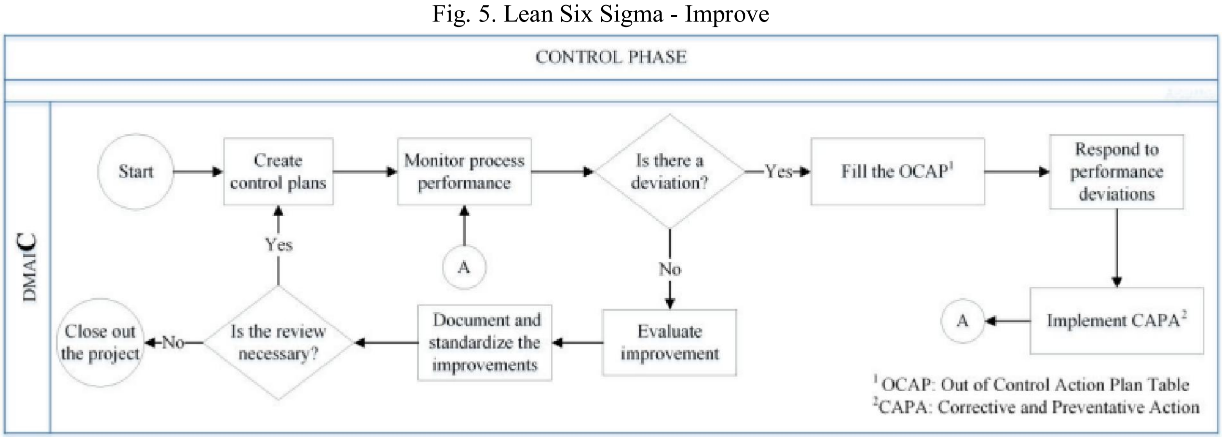

Fig. 6. Lean Six Sigma – Control

During the Control phase, tools like Failure Mode and Effects Analysis (FMEA) are utilized to mitigate potential risks, while control diagrams help assess whether the process remains within acceptable limits. Employing unsupervised

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 4, April 2024

learning algorithms derived from machine learning techniques can facilitate the creation of an alert system to detect when the process deviates from its desired state. Additionally, conformance checking ensures alignment between the created model from event logs and the actual process. Figure 6 illustrates the suggested roadmap for the Control phase.

During the Design phase aimed at creating a new process, methodologies like Quality Function Deployment (QFD) can be instrumental in incorporating customer requirements into the design. Descriptive and predictive statistical methods provide essential insights into the foundational aspects of the design. Additionally, employing data mining techniques such as market basket analysis and association rule mining helps identify products and content that complement each other effectively. Figure 7 outlines the suggested roadmap for the Design phase.

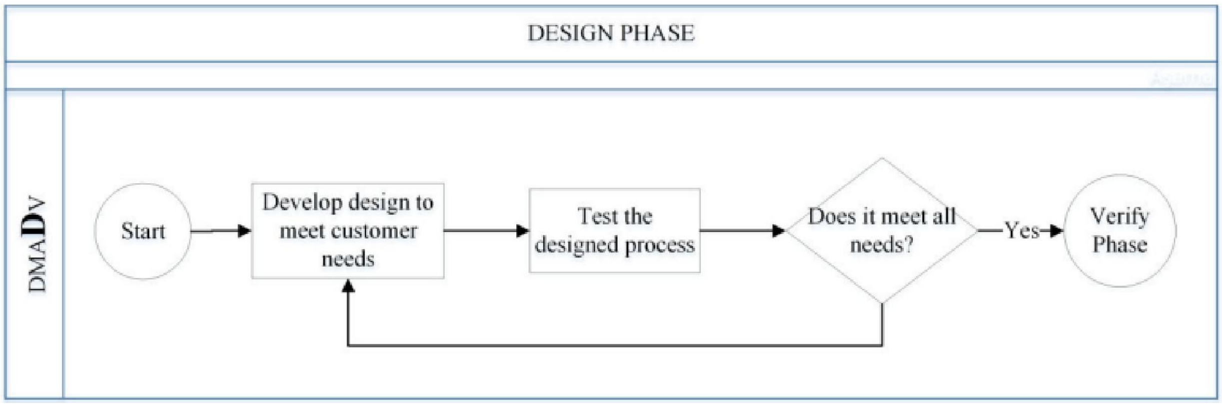

Fig. 7. Lean Six Sigma – Design

In the final step of the DMADV cycle, various metrics are established to monitor the new process, and a pilot run is conducted to validate its effectiveness. Visualization techniques such as graphing are employed to visualize results and identify any potential failures that may impact quality. Like the Control phase, a suitable machine learning algorithm can be implemented to provide alerts in case the process goes out of control. Figure 8 illustrates the suggested roadmap for the Verify phase.

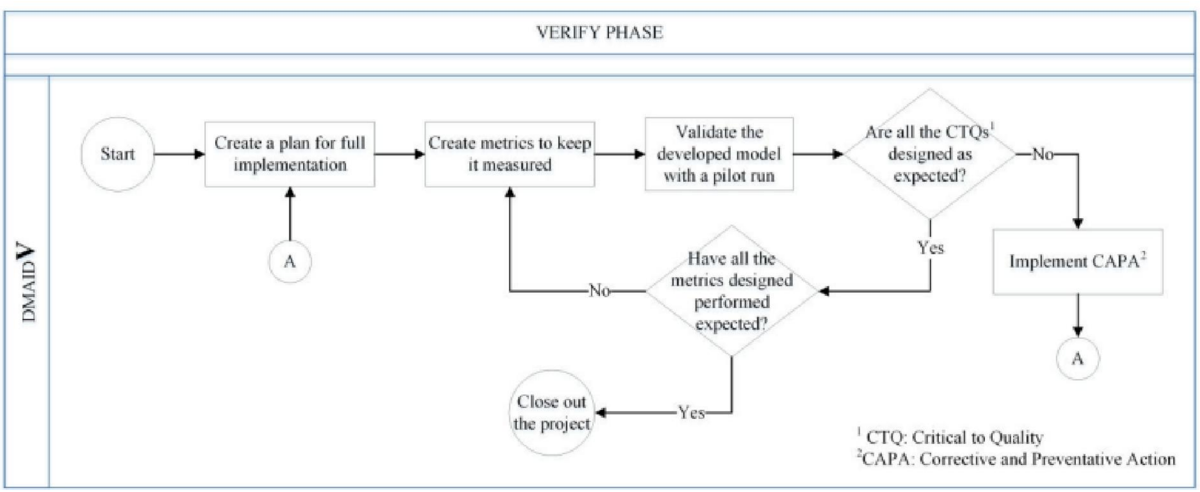

Fig. 8. Lean Six Sigma – Verify

IV. CONCLUSION

In the realm of quality improvement, several methods exist, including inspection, statistical process control, total quality control, zero defects, kaizen, and Lean Six Sigma. For this study, Lean Six Sigma is chosen as the primary method for quality enhancement, focusing on reducing variations and waste within processes through various techniques. Lean Six Sigma relies on data collection to achieve its objectives, with the collected data requiring analysis for optimal and accurate decision-making.

However, with the advent of Industry 4.0 technologies, the ability to collect vast amounts of data has become feasible. Traditional data analysis techniques may prove inadequate due to the time and cost involved. Therefore, there is a need to incorporate advanced techniques, such as big data analytics and process mining, alongside traditional methods to effectively address quality issues.

Each step of the Lean Six Sigma cycles, namely DMAIC (Define-Measure-Analyze-Improve-Control) and DMADV (Define-Measure-Analyze-Design-Verify), involves a variety of techniques, as outlined in Table 2, with separate flowcharts illustrating each phase.

The objective of this study is to provide a guide that streamlines the process, enabling faster, more reliable, and satisfactory decision-making based on data for quality improvement initiatives. As a future research endeavor, the proposed guide will be implemented in a manufacturing project aimed at resolving quality issues using the Lean Six Sigma methodology.

REFERENCES

[1]. Polk, M. (2011). Lean Six Sigma for Supply Chain Management. McGraw-Hill Education.

[2]. Guarraia, P., Hamilton, H., & Lee, P. (2008). Six Sigma: It's Not Working in R&D. Research-Technology Management, 51(2), 53-58.

[3]. Köksal, G., Uğurlu, M., & Gümüşkaya, H. (2011). Data mining applications in manufacturing processes: a literature review. Journal of Intelligent Manufacturing, 22(6), 839-850.

[4]. Van der Aalst, W. M. (2016). Process Mining: Data Science in Action. Springer.

[5]. Günther, C. W., Knoll, N., & Wölfle, R. (2017). Big Data in the Context of Industry 4.0. In Big Data in Organizations and the Role of Human Resource Management (pp. 19-36). Springer.

[6]. Rahman, M. S., & Islam, M. A. (2014). Clustering Techniques and its Applications in Health Sector. International Journal of Computer Applications, 103(14), 18-25.

[7]. Van der Aalst, W. M., Weijters, A. J., & Maruster, L. (2006). Workflow mining: discovering process models from event logs. IEEE Transactions on Knowledge and Data Engineering, 16(9), 1128-1142.

[8]. Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified Data Processing on Large Clusters. Communications of the ACM, 51(1), 107-113.

[9]. Yazdizadeh, R., & Ameri, P. (2015). A systematic review on text mining process and techniques. Journal of Intelligent Information Systems, 45(1), 1-23.