GRIFFITH COLLEGE DUBLIN

Innopharma education

GRIFFITH COLLEGE

UTILIZING ARTIFICIAL INTELLIGENCE (MACHINE LEARNING ALGORITHMS) FOR PROCESS OPTIMIZATION IN PHARMACEUTICAL MANUFACTURING PROCESSES.

By

ARAFAT ADELODUN JOHNSON

[redacted]

A thesis submitted in partial fulfilment of the requirements for MSc in Digital Transformation (Life Science)

Innopharma Education, Faculty of Science

Griffith College Dublin, Ireland

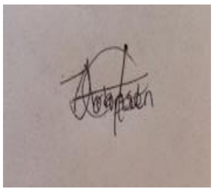
May 2024

GRIFFITH COLLEGE DUBLIN

CANDIDATE DECLARATION

I declare that the dissertation title "Utilizing Artificial Intelligence (Machine learning algorithms) for Process Optimization in Pharmaceutical Manufacturing Processes" submitted by me to Griffith College has not previously formed the basis of another research or award of any degree in any other institution. The results presented in this research are of my own work and the works of others were cited and acknowledged in areas where utilised.

CANDIDATE: Arafat Johnson
SIGNATURE:



SUPERVISOR: Sebatian Clerkin
DATE: 19th May 2024.

GRIFFITH COLLEGE DUBLIN

ACKNOWLEDGEMENT

I am grateful to God Almighty for the grace, patience, and strength he has blessed me with from the first semester to the completion of the final semester of my program in Griffith College.

I would like to acknowledge my thesis supervisor, Sebastian Clerkin for guidance and support all through the period of my dissertation and to the participants that willingly took part in making this research a successful one. Also, many thanks and appreciation goes to the dissertation committee for organizing workshops that was helpful during the course and the faculty of Innopharma Education.

I am also grateful for the continuous encouragement and emotional support that I received from my parents, siblings, and friends.

GRIFFITH COLLEGE DUBLIN

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AI	Artificial intelligence
ML	Machine Learning
DL	Deep Learning
ANN	Artificial Neural Network
PAT	Process Analytical Technology
QbD	Quality by Design
SVM	Support Vector Machine
RF	Random Forest
IoT	Internet of Things
CM	Continuous Manufacturing
DNN	Deep Neural Network
PCA	Principal Component Analysis
PLS	Partial Least Squares
PLC	Programmable Logic Controller
CPP	Critical Process Parameters
CQA	Critical Quality Attributes
FNN	Fuzzy Neural Network
CNN	Convoluted Neural Network
OSD	Oral Solid Dose
SSD	Semi-Solid Dose
DoE	Design of Experiment
NLP	Natural Processing Language
MVDA	Multivariate Data Analysis
CART	Classification and Regression Tree
RSM	Response Surface Methodology
KPIs	Key Performance Indicators

GRIFFITH COLLEGE DUBLIN

ABSTRACT

Digital transformation has introduced smart manufacturing, artificial intelligence, IoT, and advanced computerization to the pharmaceutical industry to drive Process Optimization. This plays a crucial role in the pharmaceutical industry as the complexity of manufacturing processes presents multidimensionality of product design, process development and product manufacturing data. While statistical techniques such as multivariate data analysis has made significant contribution to the pharmaceutical sector, its application can only be subjected to one process at a time in terms of providing support for quality-by-design based development and manufacturing of pharmaceuticals, limiting the enormous potential for automation. By leveraging machine learning, manufacturing processes can be streamlined to mitigate challenges associated with variability and complexity through predictive analysis of the large volume of data generated by PAT. This paper aims to provide a critical overview of how ML can be applied during various stages of the manufacturing process through a comprehensive analysis of existing literature from peer-reviewed journals, books, academic papers with illustrative examples applied in the context of pharmaceutical formulation development and related technologies as well as future trends. The study also aims to gain objective insights regarding the use of ML in pharmaceutical dosage manufacturing by exploring the opinions and perspectives of professionals actively involved in pharmaceutical manufacturing processes. With an estimated sample size of 90 participants, the study utilised an online survey-questionnaire that was administered to process managers, operators, industry experts, quality assurance and control officers to gather quantitative data in Ireland. An overall response rate of 69% was obtained and their opinion was evaluated in line with reviewed literature. The outcome of the study demonstrated the potential benefits that ML had to offer the pharmaceutical industry, the current applications, the limitations, and regulatory issues surrounding the adoption of ML in pharmaceutical manufacturing from both primary and secondary data sources.

Keywords: Machine learning, Artificial intelligence, PAT, Pharmaceutical Manufacturing, Optimization, Digitalization, Automation.

CHAPTER ONE

1. INTRODUCTION

1.1. Purpose and Significance of Study

Pharmaceutical manufacturing involves numerous complex processes that can benefit from overall process optimization and control through the model utilization of machine learning algorithms by analysing critical process parameters based on existing data and knowledge. This AI-driven approach has the potential to move the pharmaceutical industry towards a paradigm of development that is more effective. The importance of this study is to identify key areas of manufacturing that ML could potentially improve in the pharmaceutical sector by evaluating the knowledge and awareness of operators and production managers regarding the processes involved in manufacturing in the pharmaceutical facilities.

1.2. Background of study

There is a growing demand for the development of pharmaceutical products with a strong dependence on advanced computational techniques. The pharmaceutical sector is currently experiencing significant transformation through the implementation of digitization, automation and the use of big data generated throughout production processes. This has resulted in a full revolution of the industry at large. In addition, the COVID-19 pandemic also emphasises the need to modernize pharmaceutical manufacturing to produce rapid and effective medications. In recent years, numerous advanced technologies have been introduced to improve pharmaceutical manufacturing. However, there is still progress needed to establish a dynamic and adaptable pharmaceutical manufacturing industry capable of consistently producing top-quality drugs (Rantanen and Khinast, 2015).

Due to the growing need for cost-effective production with regulatory requirements and increased complexity of therapeutic molecules, there is a demand for robust pharmaceutical manufacturing processes (Dong *et al.*, 2023). With the introduction of the fourth industrial revolution (Industry 4.0), pharmaceutical devices are equipped with sensors, process analytical technology (PAT) tools and network connectivity to enable communication, data collection and sharing. In this research, machine learning algorithms, an aspect of artificial intelligence is proposed as a powerful tool with enormous potentials for supporting the optimization, and control of pharmaceutical manufacturing systems to improve product quality. Innovation in the pharmaceutical sector is usually based on research and development in a wide range of areas especially in manufacturing technology (Vora *et al.*, 2023). Providing therapeutic molecules that are suitable for use in the healthcare industry and deliver the greatest possible benefit is the main objective of pharmaceutical industries. There has been a rising focus on enhancing the safety and efficacy of drugs while reducing production cost of pharmaceuticals through the adoption of more organized pharmaceutical development and manufacturing methods. An example is the introduction of the process analytical technology (PAT) framework by the United Staes Food and Drug Administration (US FDA) for real-time process control and quality assurance and quality by design (QbD) initiatives enacted by the international conference on harmonization (ICH) (Rantanen and Khinast, 2015). The technologies of industry 4.0 has also prompted a massive increase in the number of sensors and data collected for various manufacturing processes, leading to the dramatic increase in the quantity and complexity of data stored and transferred. Machine learning has the potential to significantly impact manufacturing process analysis and improvements, demonstrating successful outcomes across several pharmaceutical industrial sectors. However, the challenges associated with this technology and procedures are dependent on data type gathered by the industry (Almanei *et al.*, 2021).

There are series of processing steps involved in parenteral and biopharmaceutical manufacturing, all of which require real-time monitoring of numerous variables. Although machine learning has made remarkable advancements in these areas, its current applications are limited in terms of providing support for quality-by-design (QbD) based manufacturing. This limits the potential for automating bioprocesses throughout the product lifecycle in the manufacturing phase.

Given the multitude of variables involved in the formulation and the manufacturing processes, pharmaceutical formulation and its associated procedures are highly complicated. Proficiency in statistical techniques including design of experiment (DoE), optimisation, and multivariate data analysis, is crucial for examining the multi-factorial relationship and variable interaction to development a product with appropriate quality attributes. The incorporation of advanced technologies such as artificial and manufacturing intelligence, in-process control and automation, digitalization, and cloud architecture into the development of pharmaceuticals daily, right through the manufacturing lifecycle, can help resolve process variabilities of starting raw materials (O'Mahony *et al.*, 2022). The implementation of technologies would also result in a significant increase in productivity, while simultaneously preserving manufacturer's competitive advantage and lowering the cost of product (Iacovelli *et al.*, 2018) (Reklaitis *et al.*, 2010). In the pharmaceutical and biopharmaceutical manufacturing, artificial intelligence (machine learning) presents a significant number of opportunities for optimising the design and control processes involved. The increasing demand for pharmaceuticals on a global scale and transition towards industry 4,0 are the major driving forces for the application of machine learning algorithms. These forces have also prompted the development of integrated process platforms and continuous processes that call for smart manufacturing and automated supervision (Rathore *et al.*, 2023). One of the primary concerns of the pharmaceutical industry in material processes is to have reliable and optimised process parameters that ensure product quality, lower costs, and higher production rates. Here, the role of smart machine learning technologies for optimising these parameters have been utilised to prevent expensive cycles of trial and error. Utilising ML-assisted procedures to optimise parameters, suggested hybrid physical-data driven technique would employ intricate analytical strategies (Horr, 2022).

1.3.Complexity and Intricacy of Pharmaceutical Manufacturing Processes.

Due to the recent advancement in pharmaceutical technology and distinctive advances, the level of complexity in manufacturing has increased in recent years. The nature of a drug product greatly influences the level of complexity in pharmaceutical manufacturing (Sarkis *et al.*, 2021). There are two (2) main categories of pharmaceutical products: small molecules (solid dosage APIs) and Biologics (vaccines, monoclonal antibodies, advanced therapy medicinal products), however manufacturing biologic drugs is more complex compared to small molecule pharmaceuticals. The first category describes pharmaceuticals that are chemically synthesized. While the second describes drug products that contain components derived from living organism. These products are distinguished by formulations that are the primary determinants of decisions related to the manufacturing process. Generic versions of small molecules are mass-produced on an industrial scale, contrarily, cell-based manufacturing of biologics involves batch or semi-batch processes that incorporates complex upstream and downstream processes, and this is a major challenge when it comes to optimising and scaling up unit operations.

For oral solid dosage (OSD) forms, different manufacturing processes with varying costs and level of complexities are employed. Direct compression (DC) is a straightforward and most cost-effective approach in contrast to more intricate procedures such as wet granulation and dry granulation (Schaller *et al.*, 2019). OSD forms are administered to achieve desired therapeutic effect and the formulation includes active pharmaceutical ingredients and excipients which can be milled, dried, encapsulated, blended, granulated or tableted (Sohail Arshad *et al.*, 2021; Nagy *et al.*, 2022). These oral solid-dose forms exist in different physical forms such as capsules, tablets, dry-powders, granules, and lozenges. Tablets are the most extensively explored aspect of solid oral dosage forms. It is comprised of a unit dose that can either be modified released or immediate released, depending on the formulation (Sohail Arshad *et al.*, 2021). The quality of the material that is processed is particularly sensitive to the powder density of the material that is entering the unit operation as well as to variations in the density that occur throughout the processing of the material. This is because many pharmaceutical unit operations on OSD forms are based on volumetric operations (Stranzinger *et al.*, 2021).

Semi-solid dosage (SSD) forms are available as a wide range of dosage forms, each having unique characteristics. In most cases, topical SSD forms are produced in the form of creams, gel, ointments, or pastes (Shukla, 2017). These formulations contain active ingredients that have been dissolved in a suitable base, as well as excipients that are suitable such emulsifiers, stabilizing agent, or antimicrobial agents. The fundamental components of semisolid dosage forms are specific to its composition. In developing the formulation, the selection of appropriate raw materials is determined by the requirements for drug delivery and specific requirements to adequate emollience and quasi-medicinal properties. SSD forms, in general, are a composition of complicated formulations that contain a variety of complex structural components (Mishra, 2018). It is essential to have a procedure that can be comprehensive and regulated, although preserved topical SSD forms do not require stringent process optimization involved in the sterile manufacturing, for example, emulsions are inherently thermodynamically unstable and thus challenging to process. Utilising manufacturing vessels equipped with programmable logic controllers (PLCs) optimizes the precision and dependability of regulating the mixing speed, flow rate and duration (Carleton, 2013).

Pharmaceutical liquid dosage forms such as syrups, emulsions, solutions, and powders for suspension refer to liquid ingestible solutions that can be topically applied, or administered intravenously (Karde *et al.*, 2023). A combination of APIs and excipients make up liquid dosage forms and provides optimal therapeutic response in a specific population with rapid onset of action upon consumption. The pharmaceutical industry has made extensive use of liquid dosage forms with their applications spanning from oral preparations and injectables to optic, nasal, rectal, and vaginal formulations. Liquid dosage forms have been widely utilised in the industry (Awad *et al.*, 2021). Modified-release liquid dosage forms are essential doe maintenance od frug stability over time which reduces the need for frequent administration and improves patient compliance.

Parenteral drug preparation involves APIs and suitable excipients that can be administered intravenously, intramuscularly, or subcutaneously as liquid solutions, suspension, or as a solid dosage form. Since biomolecules, peptides, and proteins cannot be easily supplied by any other route due to bioavailability and stability, the introduction of biotechnology has raised the demand for the parenteral route of delivery of biologics. However, parenteral drug preparation is surrounded with challenges relating to stability, solubility and manufacturability (Muheem *et al.*, 2020). Obtaining accurate prediction of the formulation and physiochemical properties with the use of A.I could potentially enhance the manufacturing process by optimizing the pH, stability, and solubility of the formulation components.

According to Tulsyan *et al.*, (2018), biopharmaceutical manufacturing involves in-depth essential planning of various process parameters for real-time monitoring of variables simultaneously. This is attributed to the variability of starting raw materials and other disruptions that occurs in the upstream and downstream processes. The purpose of conducting an analysis of bioprocess data is to gain significant insights in vast amount of data associated with numerous process iterations. These insights may be implemented to optimise and streamline the effectiveness of biopharmaceutical manufacturing procedures. During upstream processing (initial phases of cell cultivation and harvesting of APIs), identifying key process variables that greatly influences the quality attributes of the finished product and is important, hence utilizing machine learning algorithms is advantageous for assessing and comprehending the data generated during bioproduction. ML-based approaches are being used more frequently to develop accurate and flexible predictive models for analysing, monitoring, and controlling bioproduct design, bioprocess development and product manufacturing data due to their complexity and their complexity and multidimensionality in the biopharmaceutical industry.

1.4. The Concept of AI (Machine Learning Algorithms) in Pharmaceutical Manufacturing Processes.

According to a book by Gibson (2013), the concept of artificial intelligence was first adopted in the pharmaceutical sector in the early 1990s and has gained significance use in the industry ever since. McCarthy *et al.,* proposed the theory of AI during a conference in 1956 stating that “Artificial Intelligence is a process that simulates human intelligence using computers”. The approach to artificial intelligence incorporates the different fields of material science, chemical engineering, statistics, and computer science. Machine learning is an aspect of artificial intelligence that is of a specific area of research that focuses on the building of models that can learn from data, make predictions, optimization and decision making without the need for programming. In recent years, machine learning has emerged as the dominant filed within artificial intelligence in terms of its practical implementation. It is a well-established mathematical approach that reveals underlying data however ML is very broad, offering a wide range of algorithms and approaches such as supervised learning, unsupervised learning, and reinforcement learning (Selvaraj *et al.*, 2022; Wuest *et al.*, 2016).

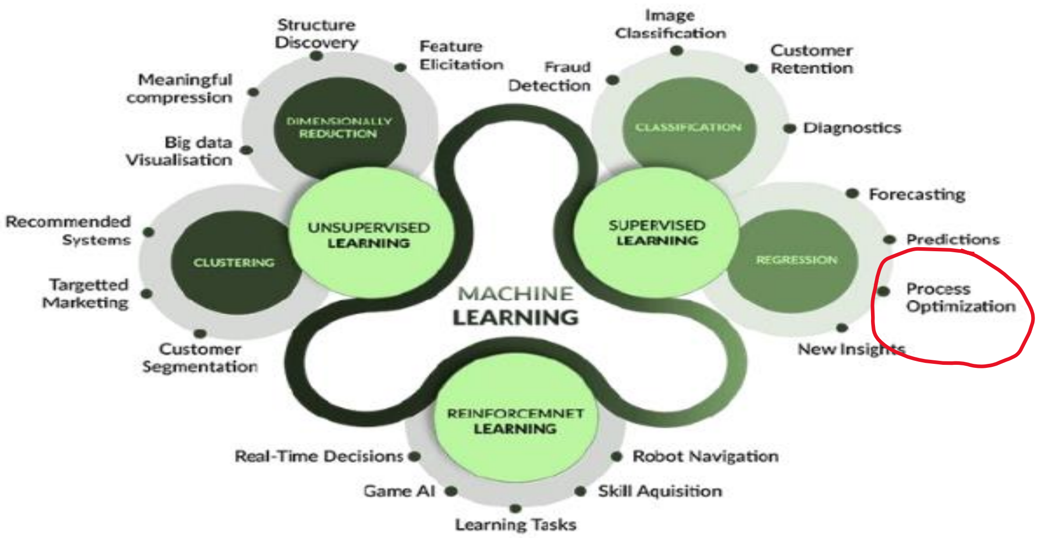


Figure 1.1 Structuring of ML Techniques and algorithms according to (Pham and Afify, 2005)

The use of machine learning in pharmaceutical sectors has grown significantly over the years due to recent advances such as deep learning (DL) by building predictive models from data (Vamathevan *et al.*, 2019). The type of the analysed data (e.g. discrete or continuous, binary, or multiple classes), as well as the type of the model (e.g. parametric or non-parametric), are also relevant for the selection of the appropriate machine learning algorithm. For instance, supervised learning encompasses techniques such as Artificial neural network (ANN) and multivariate regression and classification analysis which are designed to learn from input data and corresponding outcomes. Supervised learning approaches are often involved in process design, optimization and control (Arden *et al.*, 2021). Classification analysis sorts training data into different groups or classes based on categories. For performance evaluation, a confusion matrix is developed, with the true classes represented by rows and the predicted classes by columns. Commonly used classification methods include the Naives Bayes, k-nearest neighbour (k-NN), decision trees, support vector machines (SVM), and random forest (Munir *et al.*, 2021).

In unsupervised machine learning, inferences are drawn from input data to make predictions without relying on results to guide its learning. Dimension reduction and cluster analysis are two examples of unsupervised learning methods that can help to find operation-related trends and outliers. In the process of clustering, the objective is to identify subgroups that are comparable to one another. All the objects are categorised into predetermined number of clusters, and inputs that share a pattern brought together by the same cluster The process of reinforcement learning involves the correlation of actions with delayed results, by identifying patterns, relationships, and trends in data to understand the impact of variables on outcomes and make accurate predictions. It is often employed in complicated dynamics such as manufacturing plant operations (Ghazwani *et al.*, 2023). By utilizing the massive amount of available data, machine learning regression models can deliver significant insights and contributes to the advancement in in different manufacturing areas. Although all these ML methods could improve pharmaceutical manufacturing processes, supervised learning methods are often considered to have a lower risk and uncertainty and has been the most successful in terms of gaining momentum in the industry thus far (Arden *et al.*, 2021). By employing database technology and statistical analysis, these methods extract information and identify correlation among diverse datasets.

1.5. Challenges in Pharmaceutical Manufacturing and Optimization.

The pharmaceutical industry plays a pivotal role in advancing global healthcare by developing and producing safe, effective, and high-quality medications to treat various medical conditions. However, the process of manufacturing pharmaceuticals is complex, involving numerous stages, stringent regulatory requirements, and strict quality standards to ensure product safety and efficacy. The development and application of manufacturing techniques currently poses a variety of non-trivial challenges in the pharmaceutical sector (Rogers and Ierapetritou, 2014). These challenges arise from the complexities involved in the modelling and optimising dosage-based processes. One major challenge of implementing machine learning in pharmaceutical manufacturing is the fact that the development of model is mostly focused on technical factors and there is limited research that investigates the problems that arise from the interaction between people, process, and technology.

The pharmaceutical industry is currently facing significant challenges in developing formulations that can be directly compressed and contain high doses APIs with poor flowability and compressibility properties (Djuris *et al.*, 2021). With high API load, it is difficult to design a reliable DC formulation. Machine learning is a computational approach that connects essential flows, compaction reactions and critical quality attributes (CQAs) of basic material attributes and process parameters. This enables optimization within the design space as well as feasibility evaluation. Manufacturing biopharmaceuticals presents a number of challenges pertaining to data management and analysis as well as understanding the entire bioprocesses (Khuat *et al.*, 2024) and the include unknown impact of multidimensional changes in process parameters and development of new biologics.

The close resemblance of variants to the target protein(s) and the great complexity of protein is a challenge in the developments of biopharmaceuticals. In addition, a wide variety of amino acid sequences, disulfide and glycosylation structures, these variants can also include a very high number of distinct combinations (Hong *et al.*, 2018). The oxidation and/or deamidation of specific sites on a complex protein are examples of further modest changes that can occur. Although there has been some success in utilising approaches from the machine learning community as described by Severson *et al* (2015), there are potential in the development of algorithms that are tailored to handle the features that are unique to biopharmaceutical data.

Various research has specifically examined the utilisation of machine learning and regression techniques for the purpose of predicting lead-time in biopharmaceutical manufacturing. However, despite these advancements, the process of selecting features for lead-time prediction remains problematic. Another challenge is the development of accurate predictive models for upstream and downstream process. This is because of limited historical data and feedback to measure inherent variabilities involved in the process. Steinwandter *et al.* (2019) suggested that using digital transformation technologies and smart algorithms to monitor and optimise manufacturing processes can address these challenges, however integrating them into existing manufacturing systems would require significant investment in equipment, training, and infrastructure.

1.6.Research Aims and Objectives.

This research aims to review various aspects of pharmaceutical manufacturing operations that machine learning algorithms can potentially improve. The objectives are as follows:

1. To conduct a comprehensive review on the current state of pharmaceutical manufacturing and the existing applications of ML to various stages of process development in the mentioned areas above.
2. To identify the process parameters that need to be controlled to prevent non-conformances with regulatory requirements, explore potential applications of machine learning in detecting process inefficiencies and optimizing different aspect of the manufacturing processes.
3. To determine the type of machine learning algorithms that can be customized to handle the variability of starting raw materials and improve quality and consistency in these areas.
4. To explore the capabilities and opportunities of machine learning on efficiency and productivity for optimization operations and overall performance in pharmaceutical manufacturing.

1.7. Research Questions (RQs):

Below are three (3) research questions guiding the scope of this research. The raised questions are expected to be answered by the end of the study towards achieving the established objectives.

1. To what extent can the application of machine learning enhance operations of pharmaceutical manufacturing, what are the specific factors or emerging technologies that contribute to this AI-driven optimization approach and how do they differ across organizational context?
2. What are the opportunities associated with the integration of machine learning in the pharmaceutical sector, in terms of quality attributes, uniformity, yield, and compliance with regulatory standards.
3. How can machine learning be used in these manufacturing areas to assure public safety; what are the ethical and regulatory considerations.

1.8. Thesis Structure

This research comprises of a total of five (5) chapters and a concise summary of the content of each chapter and significance is outlined below:

Chapter 1 provides an overview and background of the study the purpose and significance of the study, background into holistic aspects of the study and highlights the aim, objectives, and questions of this research.

Chapter 2 presents an in-depth analysis of existing literature on the application of machine learning in pharmaceutical manufacturing and overview of process optimization methods focusing on a systematic and logical approach.

Chapter 3 describes the rationale behind the research and philosophical approach employed in this study. A review is conducted with emphasis on the research methodology, approach, data acquisition, gathering and analysis as well as the ethical considerations.

Chapter 4 is based on the discussion and findings. It provides a critical evaluation of the raw data obtained from primary research to identify correlations on collected data. A descriptive analysis is conducted with the use of data visualization aids, tables, and graphs. Based on actual findings, interpreted data and the result of the analysis, the concept on the utilization of machine learning in certain aspects of pharmaceutical manufacturing is then provided.

Chapter 5 concludes the work with a summary of the previous chapters, based on the findings of the results generated and proposes recommendations for future research. Detailed supporting information for selected concepts and chapters is also presented in the appendix.

CHAPTER TWO

2.1. Literature Review

In this chapter, the focus is narrowed down to the investigated trends and challenges that have been encountered in the model utilization of machine algorithms (ML) in manufacturing processes. To critically evaluate the extent to which artificial intelligence and machine learning is currently used in pharmaceutical manufacturing industries, relevant literatures would be evaluated and detailed to identify the key gaps where new information is to be contributed. Previous works of literature have been reported on the use of artificial intelligence, which spans widely on the use of machine learning techniques and neural network designs such as support vector machine, random forest, decision tress, and multiple linear regression (Ovuoraye *et al.*, 2023). Studies have shown how machine learning algorithms can revolutionize drug dosage formulation and design by predicting drug properties with the advent of automation, digitalization, and big data. According to a study by Chi *et al* (2009), machine learning is a powerful tool for data mining and optimization that can be employed to simulate pharmaceutical development as well as to determine and modify input settings and configuration for optimization.

In addition to discussing the conceptual framework defining the dependent and independent variables and their relationship to this thesis, the opinions of different scholars have been examined. Information from a variety of peer-reviewed journals, books, academic papers, news item, and newspapers was used for the secondary research, and it was useful in navigating the subject matter.

2.2. Overview of Pharmaceutical Process Optimization

The process of optimising pharmaceutical formulations involves selecting and mixing of raw materials to produce a formulation whose properties meet specific prerequisites (Chowdary and Shankar, 2016). Pharmaceutical manufacturing procedures have undergone significant transformation with emphasis on Quality by Design (QbD) tools and Process analytical technology (PAT) with the intention of enhancing the overall quality of the production process. Digital transformation has provided a science-based approach to ensuring product quality in Pharma which is founded on understanding production processes due to the quality errors with high potential of occurring. Integrating AI and ML may prove to be advantageous to the pharmaceutical industry. AI assisted manufacturing not only offers economic benefits, but it also has the potential to bring benefits in terms of designing manufacturing processes that are more robust.

To ensure product quality, manufacturing processes needs to be efficient, robust, and well understood. This is important for several reasons, including economic and regulatory consideration. With this in mind, the development of pharmaceutical processes is significantly impacted by the utilisation of process modelling and optimization technologies (Rogers and Ierapetritou, 2014). Recent years have seen significant shifts in pharmaceutical and biopharmaceutical manufacturing paradigms. This shift has been driven by the recognition among regulators, scientists, and industry stakeholder that obsolete technology is insufficient to generate new products. It is possible that the prescriptive analytic capabilities of machine learning approaches will give support to the efforts of manufacturing staff in selecting the best combination of parameters associated with a particular manufacturing process. This allows the augmentation and optimization of manufacturing processes.

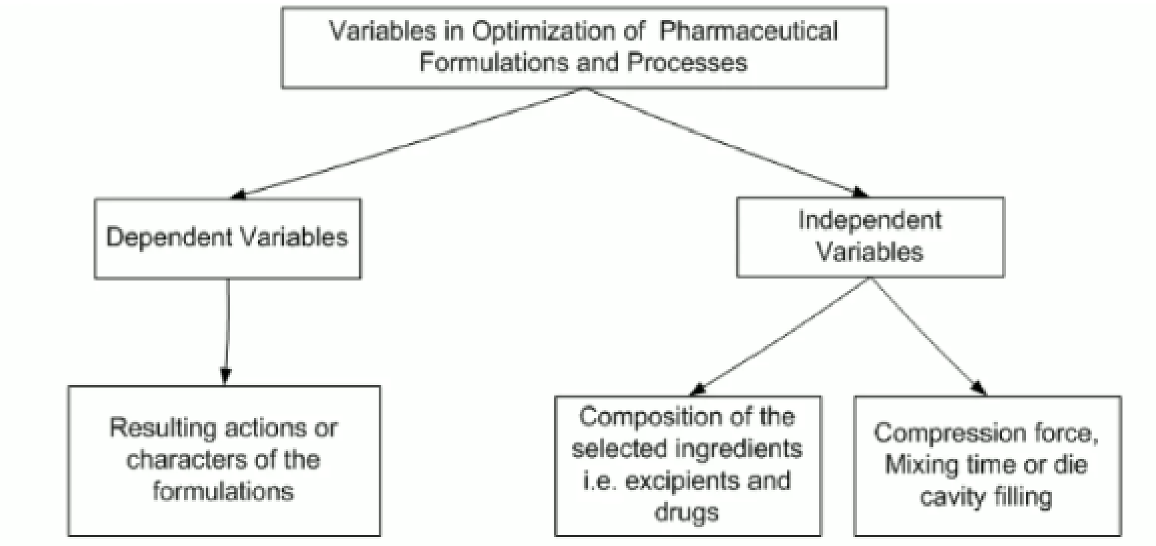


Figure 2.1 Illustration of process variables and responses involved in pharmaceutical formulation development (Beg et al., 2019).

There are several processes and manufacturing phases involved in the development of pharmaceutical products, as well as a myriad of excipients, some of which have functional properties. According to a book by Khan *et al* (2023), pharmaceutical process optimization is an area where machine learning is making significant impact through the analysis of sensors and source data to optimise processes, minimize waste and enhance energy efficiency. The process of optimizing the synthetic pathway for the synthesis of drug substances begins the selection of starting raw materials through the series of transformations that result in the intermediates of the drug substance and the final product as seen in Figure 2.2. A deep understanding of the process technology as well as the product is needed for optimization, therefore, in order to achieve the target level of product quality, a logical optimization of a number of parameters is required for product development (Beg *et al.*, 2019).

A supervised learning-based technique for maximising the application of the classification and regression tree (CART) algorithm for the line feeding mode selection is described in the paper "A supervised machine learning approach for the optimization of the assembly line feeding mode selection by Zangero *et al* (2020). The proposed approach utilises the production environment and component qualities as input to develop a decision tree, which subsequently provides a way for determining the optimal feeding procedure for each component. In addition, the study established a repair approach that presents practical alternatives with acceptable variances in average cost for scenarios that result in an unattainable solution. In term of classification, the proposed approach predicts the feeding mode with high accuracy.

In an article reviewed by Dong *et al* (2023), the traditional method of developing pharmaceuticals and manufacturing procedures might be a time-consuming and potentially hazardous endeavour. However, the industry has been compelled to move towards a more efficient development paradigm because of several global pandemics outbreaks and technologies that are associated with artificial intelligence. These technologies include chatbots, communication strategies, high-speed algorithms, and other similar technologies.

The capacity to swiftly access a wide range of data in the pharmaceutical process, as well as the efficiency of laboratories, have improved dramatically in recent years due to developments in automation technology and increasing processing power (Selekman *et al.*, 2017). This has led to the development of a number of high-throughput automation technologies that are used at different phases of the pharmaceutical process, most notably in the process of identifying routes for the synthesis of novel therapeutic compounds (Coley *et al.*, 2020). As a result of the regulations laid down by the federal agencies mandating the use of Quality by design in pharmaceutical development, design of experiments (DoE) have become an essential tool for researchers throughout the product development process. This is of great advantage to process optimization as it is the most convenient method for linking the critical quality attributes (CQAs) to the excipients as well as the process parameters (Ibrić *et al.*, 2012).

As stated in the article "Optimizing Bioprocessing techniques for Pharmaceutical Manufacturing" by Ray (2023), it is possible to produce complex and delicate biological products that cannot be synthesized through conventional procedures. This makes pharmaceutical bioprocessing an essential component of modern medicine. This technology has had a substantial impact on the advancement of numerous life-saving pharmaceutical and immunisations. However, due to the complexities involved in working with living organisms, the demand for highly settings, and the stringent regulatory standards that are imposed by the pharmaceutical sector, bioprocessing may be a procedure that is both difficult and expensive. When it comes to monitoring processes in biopharmaceutical manufacturing applications, statistical methods such as principal component analysis (PCA) and partial least squares (PLS) have been utilized for a considerable amount of time.

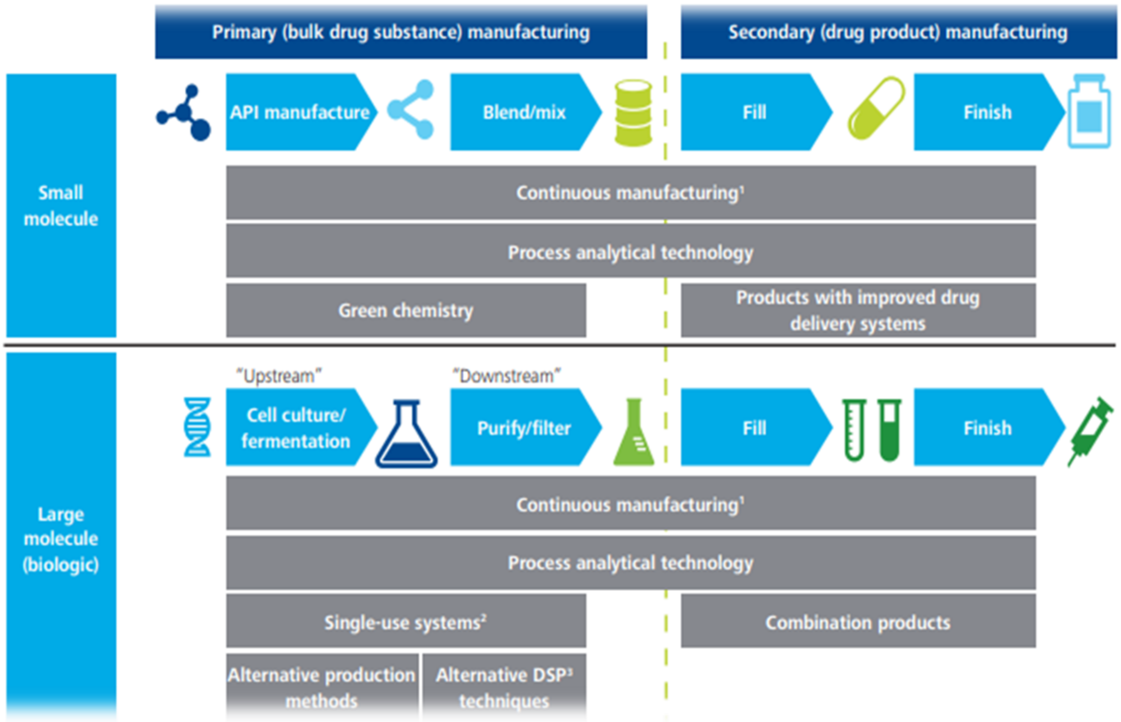


Figure 2.2. Typical process layout of pharmaceutical and biopharmaceutical manufacturing (Dong et al., 2023).

Pharmaceutical solid dosage forms have seen a major increase in interest in continuous manufacturing (CM) over the past decade. This new technology has the potential to lower manufacturing cost and waste, boost manufacturing flexibility and agility in response to variations in the demand for drug products and improve the quality and reliability of their products (Jie *et al.*, 2023). In line with the ICH Q10 and Q13 guidelines, control strategies can be implemented to keep the process under control. This means that process monitoring tools and control techniques can be used to improve product quality and ensure continued applicability and capability (Dahlgren *et al.*, 2019). There is a wide range of ML techniques that can be applied in pharmaceutical process control and optimization. Process characterization, technology transfer, clinical supply, validation, regulatory filing, and manufacturing are all facilitated by the development of the chosen path into a fully functioning manufacturing process during process optimization (Griffin *et al.*, 2023).

2.2.1 Pharmaceutical Process Optimization Techniques

The search for a practical and efficient way to optimise pharmaceutical processes is a crucial aspect of manufacturing (Kontoravdi *et al.*, 2013). In today's pharmaceutical industry, optimization has emerged as a method for best compromise solution to a problem through the utilization of bioinformatics, pharmacogenomics, and model-based technologies to expedite drug development process and curtail expenses. Optimization techniques are used where applicable in the pharmaceutical industry for process design and scaleup. Fundamental optimization of reaction processes and kinetics is common in API process development and solid oral dosage manufacturing (Vishwakarma *et al.*, 2022). These optimization methods offer a methodical framework that facilitates the discovery of favourable conditions and the application of optimization techniques leading to improvements in product development and manufacturing. By leveraging pharmacogenomics and bioinformatics, manufacturers can streamline the development process. In addition, the adoption these optimization techniques in batch and continuous pharmaceutical manufacturing can improve product quality and productivity (Escotet-Espinoza *et al.*, 2016).The significance of optimizing a pharmaceutical product is in identifying critical factors, finding a more cost effective and efficient approach to formulate the product, and enhance the uniformity and utility of quality specifications in the formulation (Chodankar and Dev, 2016; Cheng, 2023).

As stated in a review article by Shelke *et al.,* (2021), in modern pharmaceutical optimization, a methodical Design of experiment (DoE), Statistical Process Control and Process Analytical technology is employed to improve formulation inconsistencies. Traditional approaches to API and solid dosage forms such as one factor at a time (OFAT) and quality by testing (QbT) approach, are frequently used in the optimisation of analytical method and dosage forms formula. A better alternative approach is becoming necessary as the OFAT and QbT approach is inherently time consuming and chemically wasteful (Jariwala *et al.*, 2023). Pharmaceutical Formulation by Design (FbD) and (QbD) has replaced the OFAT strategy for optimization completely i.e. manufacturing changes within the approved design space without further regulatory review (Shruti, 2015).

2.3 Application of Machine Learning in Pharmaceutical Manufacturing Processes

Numerous reviews exist regarding the application of machine learning, data mining, and other AI implementation in previous years (Fahle *et al.*, 2020). For instance, Harding and kusiak (2006) provided a summary of AI applications from 1987 to 2005. The continued upward tendency in digitization only serves to reinforce the unwavering trend for machine learning. An optimal set of parameters for a certain manufacturing process can be selected with the assistance of predictive analytical capabilities of machine learning algorithms (Baviskar *et al.*, 2023). Consequently, this makes it possible to optimise and increase the production process. Effective fault detection is essential in manufacturing, especially in cases where the procedures are extensive and complex, including a substantial number of variables. It is essential to identify faults in a timely and accurate manner during production as this reduces time spent in downtime and saves energy (Rai *et al.*, 2021). Application of machine learning techniques such as artificial neural networks and fuzzy neural approaches can be utilised for the purpose of defect detection. In the pharmaceutical lyophilization process, Colucci *et al.,* (2021) utilized machine learning for the purpose of remotely monitoring the quality of the product and identifying any failures that occurred during the procedure. To identify any flaws in real time and monitor product quality, machine learning algorithms were developed with the help of data obtained from a non-invasive image sensor. Comparative research was conducted on principal component analysis (PCA) and partial least squares regression, both of which was developed together. To train the process reference model, five (5) batches were gathered in their study under standard operating conditions. A further evaluation of the categorization skills of the algorithm were carried out by employing five more batches that replicated a variety of faults. This brings about a significant improvement in the overall performance of the algorithms. Managing batch variability while simultaneously reducing off-spec product is made easier with the help of the offered techniques.

Research on the use of machine learning in multiple pharmaceutical preformulating/formulation and process development is increasingly common in both academia and the industrial sector (Lou *et al.*, 2021). The application of ML techniques ranges from drug development to pharmaceutical pre-formulation. Wu *et al* (2021) revealed in a recent article, in the early stages of drug development for molecular design and selection, ML was applied to QSAR (quantitative structure-activity relationships), a technique that correlates a compound's physiochemical qualities with its biological or chemical activity. By offering resolutions to multitude of intricate challenges such as predictive maintenance, quality control, process optimization and supply chain management, machine learning technologies has brought about a paradigm shift in the pharmaceutical manufacturing sector (Khan *et al.*, 2023). It emerged as a versatile tool to tackle arising tasks with the analysis of big data from the manufacturing plant, and development of IoT and digital twins for time series forecasting to optimize the formulation process as shown in figure 2.3 (Noorain *et al.*, 2023).

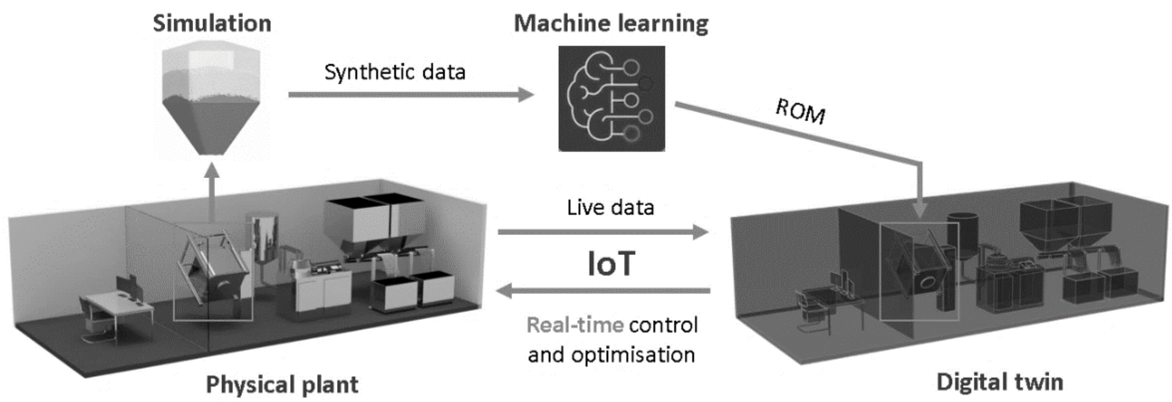


Figure 2.3 Machine Learning enhancing Automation and Digitalization in Pharma (Fischer et al., 2023)

In 2021, Wang *et al* published a detailed overview of computational pharmaceutics and Pharma 4.0. This study reviewed the perspectives of various machine learning, process simulation, mathematical models, molecular modelling and physiologically based pharmacokinetic (PBPK) modelling. The application of machine learning and deep learning is prevalent in predicting pharmacokinetic parameters (Vora *et al.*, 2023). Additionally, it can be used to examine the correlation between drug exposure and response, while considering potential confounding factors. The use of these models as shown in table 2.1 can optimize dosage regimen for a selected study/population where data is limited and (Liu *et al.*, 2021). In addition, Wang *et al* provided a summary of the regulatory requirements, problems, and prospects in that aspect.

Table 2.1. Compilation of frequently investigated ML models in Pharmaceutical Product Development.

Machine Learning Models	Description/Usage	References
Support Vector Machines (SVM)	In dosage form optimization, SVMs have been employed to predict and simulate the interaction between formulation factors, including drug release profiles, processing parameters, and excipient composition. They help make formulation design space more optimal	(Mukhamediev *et al.*, 2022)
Artificial Neural Networks (ANNs)	Drug release kinetics from various dosage forms have been optimised using artificial neural networks. The best formulation can be found as thy help predict how APIs are released under different processing circumstances	(Sun *et al.*, 2013)
Genetic Algorithms	Natural selection and genetics are the guiding concepts of genetic algorithms, which are optimization methods. Utilising them can help get the intended dosage form characteristics by optimising drug release profiles, formulation compositions and process factors.	(Bannigan *et al.*, 2021)
Artificial Neural Network-Genetic Algorithm (ANN-GA) Hybrid Models	Dosage form optimisation has been done fusing ANNs and GA methods. They can predict formulation properties and finding optimal solutions by effectively searching the formulation space	(Vora *et al.*, 2023)
Multivariate Analysis Techniques	Dosage form optimization has made use of multivariate analysis techniques including principal component analysis (PCA) and partial least squares (PLS). They support the reduction of dimensionality, identification of crucial formulation variables, and the optimisation of formulation performance	(Shi *et al.*, 2021)
Response Surface Methodology (RSM)	By analysing the relationship between several variables and their impact on formulation responses, RSM is a statistical technique that aids in the optimization of dosage formulation. This facilitates comprehending and maximising formulation parameters.	(Prasada *et al.*, 2020)

With machine learning approaches, it is possible to discover optimal manufacturing strategies for a given material and provide rapid digital screening tools for advanced pharmaceutical development by implementing machine learning models in the early stages of drug development. Even though supervised learning and unsupervised learning approaches have already been widely utilised in pharmaceutical manufacturing, accounting for about 90-95% of all applications, reinforcement learning has been subjected to a lesser amount of research. As a result of this, the following part provides a selection of the most significant research roadmaps for supervised and unsupervised learning in the pharmaceutical industry (Gómez Escudero *et al.*, 2021).

The ability of machine learning algorithms to uncover previously (hidden) information and implicit connections in data sets is one of the most significant advantages to using these algorithms, as was mentioned before. Different aspects of ML approaches, such as supervised or unsupervised learning or reinforcement learning, can result in different requirements for the data that is provided. The broad capacity of machine learning to generate outcomes in a manufacturing setting has been demonstrated by several studies (Mohammed, 2022).

2.3.1. ML in API and Solid Dosage Manufacturing

In the aspect of API manufacturing, Strachon and Schongut, (2023) reviewed the application of machine learning for predicting the bulk behaviour and particle properties of active pharmaceutical ingredients. Based on the fundamental particle attributes of APIs such as particle size and shape distributions, the purpose of their research was to develop predictive models using random forest regressor and multi-layer perceptron (MLP) regressor for powder behaviour by flow function coefficient for formulation and process development. The hyper parameters of these models were evaluated using leave-one-out cross-validation and the results showed that the models could predict the flow function coefficient with moderate to high accuracy. The research uncovered both the potential and the constraints of machine learning to predict the bulk behaviour of powder.

In solid dosage manufacturing, numerous ML-based model has successfully been applied in pharmaceutical solid dosage form development in recent years. These formulations have made substantial use of many supervised learning techniques including decision trees, linear regression, logistics regression, random forest, XGBoost, lightGBM and support vector machine. A recent journal article by Diaz *et al.*, (2023) describes the current use of machine learning approaches to the prediction of powder flow behaviour of pharmaceutical materials to physical properties in solid dose manufacturing. According to their study, machine learning models have the capacity to expedite decision making and minimize time and materials needed to develop robust processes in pharmaceutical manufacturing. Their study focused on the application of ML models to predict flow properties of common, widely accessible pharmaceutical materials. These models were trained using 10-fold cross validation to ensure performance and evaluated on the physical parameters in terms of bulk density, size, and shape of pharmaceutical powders.

To gain a deeper comprehension of how ML algorithms can be utilised in various solid dosage formulations, table 1 provides a comprehensive overview of recent ML algorithms in this aspect. This summary provides a comprehensive and complete understanding of the research conducted in this area. The number of articles that are relevant to ML in solid dosage forms has increased by 100% on an annual basis from the year 2015 as stated in the following published literature (Jiang *et al.*, 2022).

Table 2.2. Overview of several applications of ML in oral-solid dosage forms.

Dosage forms	**Applications**	**Algorithms**	**References**
	Detecting tablet defect.	CNN, XRCT (computed topography)	(Ficzere *et al.*, 2022) (Ma *et al.*, 2020b)
	Tablet formulation using 3D printing technology.	ANN, self-organizing maps, RF, SVM and CNN	(Westphal and Seitz, 2021) (Muñiz Castro *et al.*, 2021)
	Analysis of disintegration rate.	RF, XGBoost, ANN, and CNN	(Szlęk *et al.*, 2022) (Floryanzia *et al.*, 2022)
	Predicting drug release	ANN, SVM, decision tree, and ensemble of regression tree	(Galata *et al.*, 2021) (Salem *et al.*, 2022)
Tablets	Inspection of drug particle size	Pattern recognition neural network.	(Mészáros *et al.*, 2022)

	Regulatory oversight of the powder engineering process	ANN	(Chauhan *et al.*, 2021)
	Formulation design of dry powder for inhalation	RF, XGBoost, LightGBM, SVM, KNN, ANN, and CNN	(Kazemzadeh Farizhandi *et al.*, 2021)
	Predicting dry powder particle size distribution and extent of agglomeration	SVM, RF, and partial least squares regression	(Xi *et al.*, 2020)
Powders	Optimising the compatibility of spray-dried powder.	SVM and ANN	(Lou *et al.*, 2019)
	Identifying particle defects contained within the capsules.	SVM	(Dörr and Florence, 2020)
Capsules	Identifying capsule defects	KNN, SVM, and CNN	(Zhou *et al.*, 2020)
	Predicting particle size distribution	ANN, multiple linear regression, and genetic programming	(Kazemi *et al.*, 2016)
Granules	Granulation process control	Neuro-fuzzy logic and genetic programming	(Lou *et al.,*2019)
	Predicting dissolution rates and profiles	RF, SVM, LightGBM, and XGBoost	(Dong *et al.,* 2021)
Solid dispersions	Predicting physical and chemical stabilities	ANN, SVM, RF, LightGBM, KNN, and Naives Bayes	(Lee *et al.*, 2022); (Han *et al.,* 2019)

Tablet characteristics such as binding, capping, discoloration, binding or variations in shape and size frequently occur during the manufacturing process (Jiang *et al.*, 2022). Manual screening of defective tablets requires many workers, which is quite complex during scale-up processes. To overcome this challenge, analysing the interior structure of tablet may be accomplished by the utilization of X-ray computed tomography (XRCT). Combining XRCT with deep learning has allowed researchers to effectively identify tablet defects, expanding the technique's potential application. Ma *et al* explored the application of convolutional neural network for the purpose of identifying internal defects in tablets. As part of their research, numerous batches of tablets were prepared using microcrystalline cellulose and mannitol as excipients. These tablets were then subjected to XRCT imaging for the purpose of image analysis in the field of pharmaceutics.

In the pharmaceutical industry, disintegration is a crucial quality attribute for small molecule formulations since it is the first stage in the distribution of oral solid and, which is a key phase in the process (Floryanzia *et al.*, 2022). As a result, the disintegration time is one of the critical quality attributes for orally disintegrated tablets (ODT) that is a significant variable capable of being optimised. Utilising ML to visualise tablets and asses the degree of disintegration is a promising solution to the challenges that were earlier discussed by studying the samples of available data and learning how it is effectively processed as shown in figure 2.3. A neural network was utilised by Han *et al* in 2018 to predict the disintegration time of solid dosage forms. A dataset consisting of 145 formulation records comprising 23 APIs was used to train their models and this demonstrated an accuracy of 80% of the testing datasets during this research endeavour. Molecular descriptors were incorporated to describe APIs, and the dataset was duplicated and improved by removing the errors found.

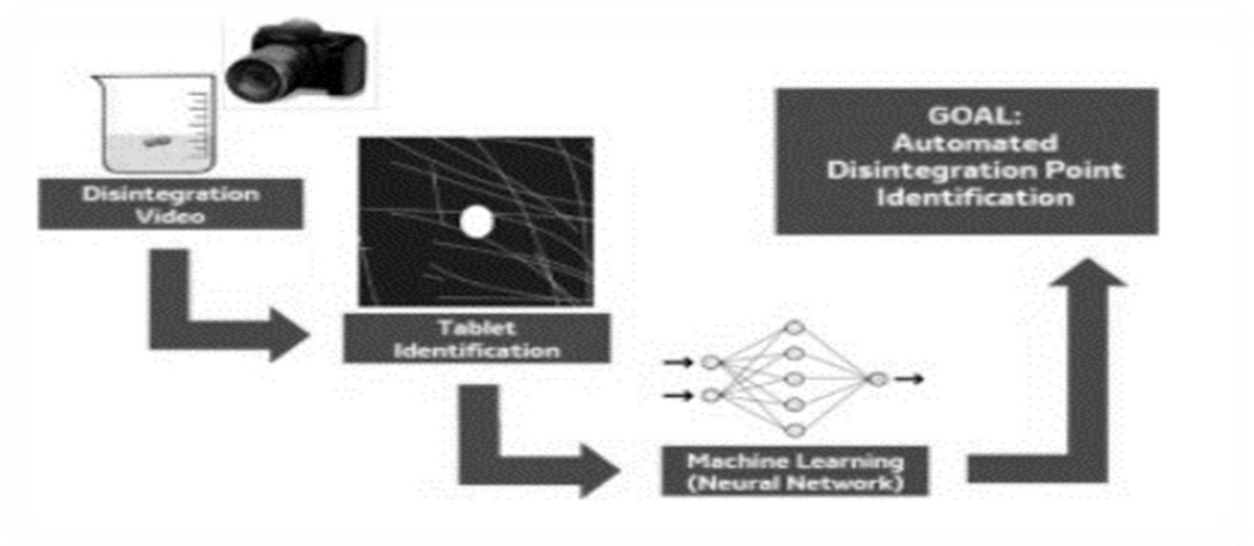


Figure 2.4. Combining Computer vision and Neural Network for detecting Tablet Defect (Floryanzia et al., 2022)

The existing literature on the utilization of machine learning in the development of capsule-based formulations is limited. Using a convoluted neural network, Zhou *et al* successfully identified capsule defects. The study utilized K-Nearest Neighbour (KNN) and SVM for the purpose of comparison. Following the implementation of this enhanced convoluted neural networ, the confusion matrix findings revealed a detection accuracy for the identification of different capsule defects. Cui and Zhu reviewed in their study how artificial intelligence makes prediction using a neural network known as ResNet of the physicochemical characteristics of various drug forms. Enhanced productivity from different origins demonstrated the superior accuracy of their framework in predicting the solubility of compounds in contrast to non-AI-based models. This illustrates the potential integration of machine learning algorithms into drug dosage forms development process to enhance the manufacturing processes.

Research conducted by Margulis, and colleagues demonstrated the application of machine learning in the field of solid-dose drug discovery. Their study focused on the early phases of formulation development and their objective was to determine whether the bitterness of different compounds employed in the medications can be predicted using a particular machine learning method, therefore serving as a viable substitute for animal testing. The study demonstrated a high level of success as 80% of the bitter compounds found corresponded to those acquired from a brief access taste aversion (BATA) experiment. The results of the BATA experiment showed that toxicity and bitterness are not always correlated, contrary to earlier assumptions. This illustrated that ML could provide accurate results as well as new information. Simoes *et al* (2020), recently published a paper on artificial neural networks ANNs and explored its use in quality by design approach, based on poorly soluble and permeable drugs of oral solid and liquid dosage forms in the pharmaceutical sector and compared the reference product with a biomedical study. In their risk assessment framework, the material attributes and process parameters that were found to be essential for the dissolution of the critical quality attributes includes tablet hardness, impeller speed, particle size distribution, sieving of the dried granules, the duration of granulation process and amount of liquid used. Multilayer perception networks MLP were trained and validated using a fully connected architecture. ANNs were then employed to establish criteria for products and processes using similarity variables for the projected dissolution profiles. Three large-scale batches with varying sets of drug samples were used in the trials to validate the best ANN model. The dissolving properties of the generated batches were precisely predicted by the ANN model. Their work showed how machine learning may be used to manufacture oral solid doses with great efficiency.

Polykovskiy *et al* conducted a study in 2018 that examined the ability of machine learning to predict the activities of small molecule APIs showing gains in efficiency. Their aim was to ascertain if machine learning might enhance the precision of the drug screening procedure, hence improving its reliability. An adversarial auto-coder was employed to predict this activity. The outcome (a compound generated via de novo design) was different from that generated based on a Recurrent Neural Network (RNN). This not only indicates that these technologies can be used in conjunction with one another but also revealed vital information target-molecule which enables procedural accuracy and efficiency.

2.3.2. ML in Semi-solid dosage Manufacturing

Research on the ML advances made in understanding semi-solid formulations and topical drug delivery is limited. However, traditional techniques cannot address variability. Taking a quality by design (QbD) approach can be difficult when dealing with semi-solid dosage forms because of their complex behaviour as some semisolids can act like a liquid and a solid within the same formulation at different times under different conditions as stated in a review article by Herbig *et al* (2023). The paper revealed information on the status of the development of semi-solid formulations, specifically highlighting contemporary technologies to examine and optimize the stability of active ingredients applicable requirements.

By leveraging machine learning, techniques, manufacturers can analyse vast amount of data and identify patterns to automate processes, optimize formulation and predict manufacturing outcomes. The use of machine learning would further aid in the optimization of stability and shelf life of semi-solid products by analysing various factors such as formulation ingredients to gain insights into the status of the development of semi-solid formulations (Sakthivel *et al.*, 2021). It is relevant to analyse the product which entered the market. SVMs and ANNs algorithms can quickly detect quality issues and provide real-time feedback, enabling manufacturers to take corrective actions and prevent defective products from reaching the market (Bharat and Paresh, 2023).

2.3.3. ML in Liquid Dosage Manufacturing.

When it comes to liquid products, homogeneity is the first of the critical manufacturing parameters to control. In the case of gels containing both organic and aqueous phases, this parameter becomes more important not only in the process control but also within its stability test (Bautista and Amigo, 2015). Artificial neural networks have been used to create stable emulsions and solutions, which are able to model and optimise the complex interactions between formulation factors and how those relationships affect overall performance and product quality (Tyagi *et al.*, 2023). In a study reviewed by Kumar *et al* in 2011, ANNs was used to construct stable oil-in-water emulsions and increase the amount of a fatty alcohol. The input data consisted of liquid paraffin, lauryl alcohol concentrations and time variables, whereas the outputs included zeta potential, viscosity, particle size and conductance. Validation experiments demonstrated the high performance between experimental data and ANN-predicted values. This is a great subject of research with practical applications in pharmaceutical manufacturing of liquid dosage formulation.

2.3.4 ML in Biopharmaceutical Manufacturing.

The biopharmaceutical sector is gradually embracing digital transformation with a strong emphasis on incorporating innovative technologies such as artificial intelligence and machine learning to fully automate their production process (Mondal *et al.*, 2023). Presently, the most profitable medication in the market are biopharmaceuticals such as monoclonal antibodies (mAbs) and fusion proteins which accounted for 70% of the to 10 medications in 2019 (Walsh *et al.*, 2022). According to a study by Guerra, biopharmaceutical manufacturing still requires continuous assessment for both the process and product. This is because various factors such as environmental disruption, gradual deviations in the process, such as fouling, loss of cell activity, inhibitors or activators and disruption in the process such as impurities in the feedstock, variations in the quality or manual sampling have the potential to impact the outcome of the process.

Machine learning algorithms can acquire insights into the dependency between CPPs and CQAs in the process of developing and producing biopharmaceutical products. This is accomplished by monitoring the interactions between variables and extracting important information from bioprocess data (Khuat *et al.*, 2024).. Further development of more effective process control strategies may be facilitated by these insights. Supervised ML algorithms is mainly used for biological activity prediction through classification and regression and can be utilised not only for the real-time monitoring and control of the upstream and downstream bioprocesses, but also for the discovery and design of biopharmaceuticals, including the selection of mAbs, media screening, feeding strategy design, and substrate selection for purification (Khuat *et al.*, 2024, Kumar *et al.*, 2022) . Figure 2 illustrates how algorithms such as support vector machine (SVM), K-nearest neighbour (KNN), Naïve Bayesian, decision tree (DT), and random forest (RF) can be effectively employed for the purpose of system training.

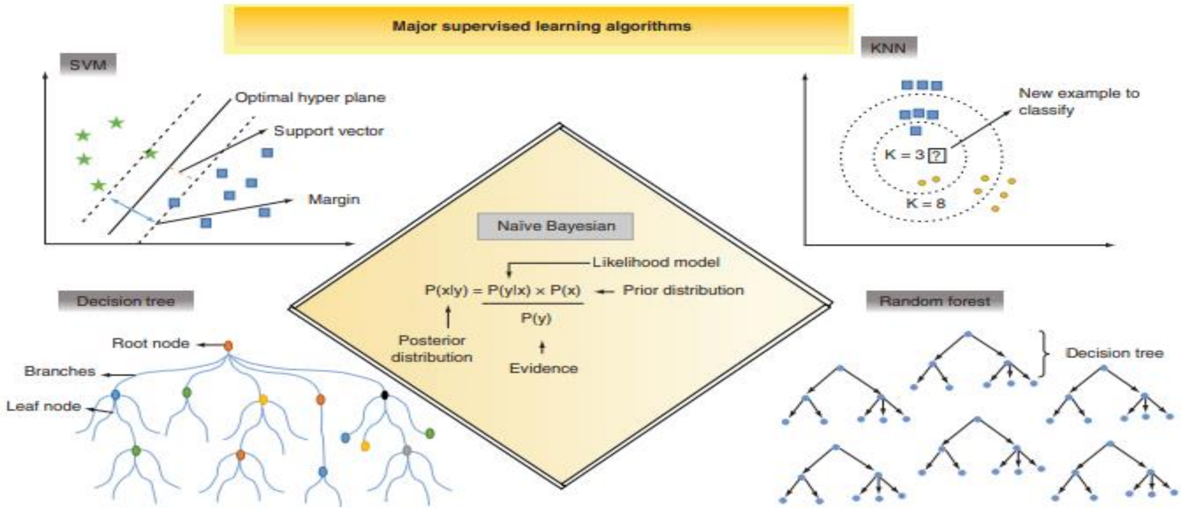


Figure 2.5. Major supervised based learning algorithms used in biopharma manufacturing (Kumar et al., 2022).

As a means of separating data into a lower-dimensional plane, the support vector machine (SVM) makes use of nonlinear kernels. This makes it easier to fit the data into a space with a maximum number of dimensions (Yang *et al.*, 2019). For pattern identification and statistical estimation, the non-parametric technique known as k-nearest neighbour (KNN) makes use of data classification. The probabilistic classification using Naïve Bayes is based on the Bayes theorem and the premise that features are independent of one another. The data can be split using DT, which is made up of nodes and decision tree. The nodes represent divisions of the data, and the leaves represent the results of separation (Reutlinger and Schneider, 2012). The set of nodes, branches, and leaves that make up the decision tress is what is referred to as the RF. To improve both training sets and predictions, the RF employs the bagging approach. As shown in figure 3, clustering and dimensionality reduction are both performed by supervised learning using neural networks, hidden Markov Models (HMM), Gaussian mixtures, Fuzzy C-means, K-Medoids, K-means, principal component analysis (PCA), and linear discriminant analysis (LDA) methods (Kumar et al., 2022).

Table 2.3 Biopharmaceutical manufacturing challenges with prospective digital solutions aimed at overcoming them (Steinwandter *et al*., 2019).

Current challenge	**Potential digital solution**
Development and approval of new biologics is costly and the time from production to clinic is too long	Increased use of digital tools for experimental design, data-driven decision-making, and bioprocess optimisation
Bioprocesses are not transparent	Using predictive soft sensors to indirectly estimate difficult or impossible-to-measure variables
Routine manufacturing processes are not cost-effective, operate sub-optimally, or lead to failed batches	Applying optimisation approaches to optimise performance of already established processes; implement enhancements for product life-cycle management
Impact of multidimensional changes in process parameters on CQAs is usually unknown	Use process modelling approaches to achieve data-driven estimates for the changes of process parameters in consideration of whole process value chain
Shortage of highly skilled workers in statistical analysis and machine learning	Deploy automated algorithms to replace the manual data analyses
It is difficult for humans to analyse and understand high-dimensional bioprocess datasets	Deploy data-driven methods for data interpretation

Currently, the distinctive characteristics and intricacies of the biopharmaceutical sector leads to diverse challenges concerning data management and analysis, as well as the understanding of the entire bioprocess. The implementation of digital transformation and application of smart algorithms for observation, management and enhancement of biopharmaceutical processes can address these challenges. Table 2.2 provides the current challenges along with prospective machine learning solutions.

As shown in Mitchel (2013), common CPPs that can impact the CQAs of the final product that are typically extracted from the cell cultures in bioreactors include pH, dissolved O2 and CO2, temperature, pressure, osmolality, nutrients, and concentration of secondary metabolites. According to Bhambure *et al.,* (2011), high-throughput cell culture media facilitates numerous testing conditions which allows for the evaluation of multiple parameters in media optimization using statistical tools as well as experimental design. For cell culture optimization, statistical and mathematical method (DoE and RSM)) are typically employed Singh *et al* (2016); Alwis *et al.,* (2007), however, it is only an effective strategy when there are less than ten components that require adjustment. Machine learning has emerged as a powerful tool to address this limitation of medium development and has been shown to outperform DoE and RSM (Grzesik and Warth, 2021)

With the utilization of CPP historical data during cell cultivation in the upstream bioreactor, ML predictors and multivariate data analysis methods has the potential to predict CQAs and key performance indicators (KPIs) (Cosenza *et al.*, 2021). Table 2.4 summarises the role of major supervised and unsupervised based learning algorithm used from strain selection to bioprocessing.

Table 2.4. Application of Machine learning algorithms during bioprocess development (Reddy, 2023).

Machine learning	**Algorithm types**	**Models/methods**	**Optimization role in biopharmaceutical manufacturing**
Unsupervised learning	Clustering	Hierarchical clustering	Deep feature selection for biomarkers
		K-means	Feature reduction in single cell data to identify cell types.
		Neural Networks (DAENS and autoencoders	Cell types and biomarkers from single-cell RNA data
Supervised learning	Classification	NLP Kernel methods	Target-disease drug associations
		NLP Bayesian Classifier	Target drug ability based on pharmaco-kinetics (pk) properties.
		SVMs	Target drugability-based protein structure or sequence.
		Nearest Neighbour	Tissue specific biomarkers from gene expression signatures

		Random forest	Molecular features that predict drug response
		Decision trees	Disease and target drug ability from multi-dimensional data
	Regression analysis	Sparse Linear Regression	Gene expression data that predict clinical trial success

Additionally, the optimization of API synthesis by ANNs has been the subject of numerous research and crystallization is crucial in the downstream formulation step to provide solid crystalline API (Nagy, D. Galata, *et al.*, 2022). For instance, using historical data, ML can estimate the crystallization outcome. According to an article by Velasco-Mejia *et al* in (2016) genetic algorithm models and ANNs was used for the optimization of industrial batch crystallization. In their study, they modelled the resulting crystal density using nine descriptors, which led to the identification of the most important factors and, following optimization, a notable improvement in the product.

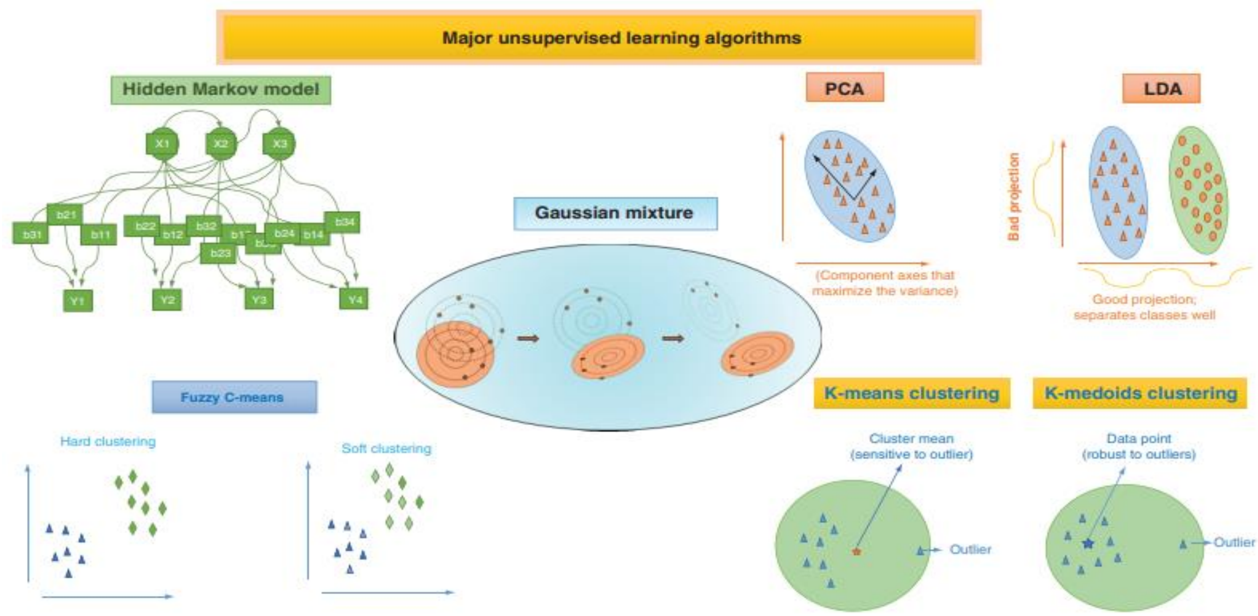


Figure 2.6. Major unsupervised-based learning algorithms in biopharma manufacturing (Kumar et al., 2022).

2.3.5 ML in Parenteral Manufacturing

Numerous studies have investigated the application of machine learning in parenteral manufacturing. According to a study by Mohan *et al* (2022), manufacturing parenteral pharmaceuticals is a challenging and a complex procedure. As the demand for sterile preparations has grown, pharmaceutical manufacturers have shifted to computerised production to minimise manual errors. Additionally, the use of ML in parenteral production is growing rapidly and aids in the detection of a variety of challenges as well as the reduction of false rejections and visual inspection of parenteral formulations. Major supervised learning (Neural networks) and unsupervised applications in this field yield more refined findings as presented in table 2.5. In comparison to other manufacturing processes, parenteral manufacturing of biologics has been slow in adopting machine learning based techniques (Kontoravdi *et al.*, 2013). Vora *et al* also stated in the journal article "Artificial Intelligence in Pharmaceutical Technology and Drug Design" the roles of ML applications in parenteral manufacturing and they are;

- ML optimises complex parenteral product manufacturing equipment maintenance operations by analysing complex data, equipment performance history and maintenance records to predict equipment failure, and schedule proactive maintenance.
- Compliance with regulatory requirement and GMP pertaining to parenteral and complicated biological product is also enhanced by ML, through the examination of process data and product attributes, machine learning algorithms may assess compliance, identify potential noncompliance issues, and offer suggestions for process improvement.
- Large datasets from analytical testing, such as particle size analysis, spectroscopy, and chromatography, may contain trends and variations in product quality that can be discovered by ML algorithms, by addressing quality issues early to ensure high-quality products.

2.4 Benefits of Using Machine Learning for Pharmaceutical Process Optimization

The basic benefits of machine learning have been demonstrated in figures and tables, where it is stated that ML algorithms may manufacturing problems, which frequently arises in the context of optimising intelligent manufacturing systems (Wuest *et al.*, 2016). However, it must be acknowledged, that the specific advantage may very based on the type of algorithm selected, for example, SVM and distributed hierarchical decision tree can handle dimensionality better than others (Bar-Or *et al.*, 2005). Several applications of machine learning techniques have proven to be useful in for general purpose needs such as natural language processing (NLP), pattern recognition and computer vision (Ailisto *et al.*, 2023). However, a recent term co-occurrence review identifies machine learning as an emerging trend in pharmaceutical manufacturing research. AI and ML have recently been associated with smart manufacturing and Industry 4.0 practices (Sharp *et al.*, 2018). Supply chain management, quality control, maintenance management, and process optimization are just a few of the uses for AI and ML that pharmaceutical sector is keen to leverage (Zheng *et al.*, 2018).

As stated by the WHO, artificial intelligence is used to access the physical and chemical characteristics of pharmaceuticals, and to ensure their quality and consistency during manufacture. This would include use in process design and monitoring, waste reduction, quality control and fault detection., with the use of machine learning for single-step and multi-step retrosynthesis (Geneva: World Health Organization, 2024). Also, the capacity of ML algorithms to handle large dimensional data is one of its advantages., given the lack of transparency in complex manufacturing data (Monostori, 2003).

Considering the unique characteristics of different processes and variables during pharmaceutical manufacturing, one of the advantages of ML algorithms provides the opportunity to learn from the dynamic system and to some extent, automate environmental adaptation (Flach, 2022). Depending on the ML algorithm, the adaptation process can be faster than the traditional approaches which serves as the foundation for predicting the manufacturing system as utilising ML techniques can lead to the extraction of patterns from existing data sets. The additional data helps manufacturer to make decisions or could be directly utilised to optimise processes by identifying specific patterns and irregularities (Pham and Afify, 2005, Monostori, 2003). Overall, it is acknowledged that ML techniques can decrease cycle time and scrap while increasing resource efficiency in the pharmaceutical industry. Additionally, machine learning offers powerful tools for ongoing quality enhancement in expansive and complex manufacturing systems. The benefits of machine learning span across process monitoring and control of drug formulation and development as well as predictive maintenance as discussed in in the following section.

2.4.1 *ML In Dosage Formulation and Development*

Multivariate tools and artificial intelligence approaches (genetic algorithms) have been observed to perform well in the context of understanding the underlying relationships between the critical quality attributes and process parameters. ML modelling methods have also been applied to the formulation development of solid dosage form, where they have observed to perform well. During product development, ANNs have been used primarily for formulation and process optimization. ANNs have also proved beneficial in the optimization of drug release rate from controlled-release formulations which are particularly complex in terms of the number of interacting factors which influence drug release (Hardy and Cook, 2010)

2.4.2 *ML for Predictive Maintenance in Pharmaceutical Manufacturing*

In pharmaceutical research, artificial intelligence has become a potent tool for predicting the physiochemical stability of solid dosage forms. AI, by utilizing machine learning methods and computational models, can predict the stability of solid dosage formulations by analysing and interpreting extensive datasets, including pharmacological variables and formulation factors. Many research endeavours have focused on the problem of predicting quality attributes for solid dosage forms. For instance, the prediction of the granule particle size distribution and tensile strengths of solid dosage dorms using regression and random forest respectively (Szlęk *et al.*, 2022).

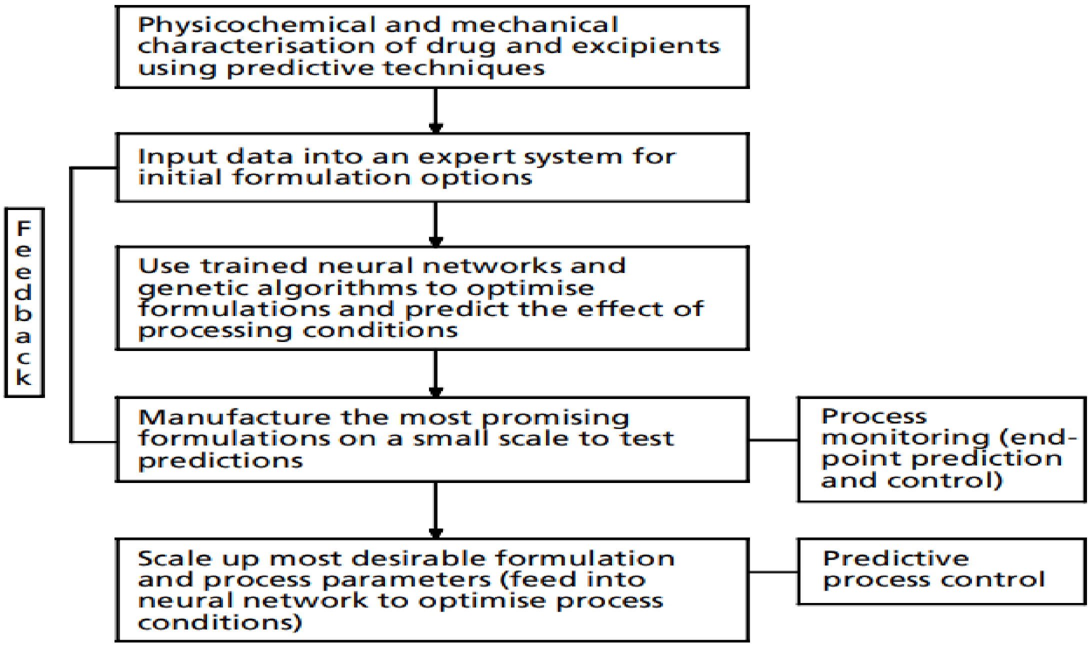


Figure 2.7 Flowchart of product development cycle incorporating ANNs predictive technique (Hardy and Cook, 2010)

Recently, numerous methods have emerged for response optimization. Although recent machine learning and artificial intelligence models for solvation and solubility have prospects for predicting appropriate green solvents and telescoping techniques, however, yield prediction remains a barrier (Griffin *et al.*, 2023). For predicting and comprehending the relationship between process parameters and response in physical, chemical and biological processes, machine learning approaches can be applied similarly to other computing techniques by extracting information, correlation and hidden patterns from various datasets using database technology, statistical analysis, and other disciplines (Liu *et al.*, 2022). For example, ANNs has been demonstrated and applied in predicting the solubility enhancement impact of hydrotrope molecules, which may be used to optimise the pre-formulation stage. In addition, the application of machine learning to predict drug release profile of minitablets, the tensile strength and physical stability has gained traction in pharmaceutical formulation processes (Barmpalexis *et al.*, 2018; Damiati *et al.*, 2017; Onuki *et al.*, 2012).

2.5. Challenges And Limitations Encountered in the Implementation of Machine Learning in Pharma

Machine learning has emerged as a promising tool in pharmaceutical manufacturing, offering the potential to improve efficiency, speed, and cost-effectiveness. However, the implementation of machine learning in pharma also comes with its fair share of challenges and limitations. A consequence of the recent advancement in computational capabilities is the heightened accessibility of commercial software, which diminishes the expertise required for implementing ML models (Banner *et al.*, 2021), and without thorough comprehension of the model's assumption, considerations, and limitations, can lead to inappropriate applications. Neural networks often pose greater challenges for interpretation due to their complex structures. This aspect holds significant importance in drug formulation and selecting new pharmaceuticals, as without proper justification could hinder regulatory approval. In such instances, it is necessary to conduct further analysis and evaluation of a model's performance and resilience before putting it into action (Mendez *et al.*, 2019).

As stated in a book by Flores-García *et al* (2023), the challenges often encountered in applying ML to pharmaceutical manufacturing includes feature selection and engineering. Generally, feature engineering and selection rely significantly on the expertise of the workforce (understanding a variable and ensuring that an approach is reproducible under different conditions). Furthermore, the selection of an algorithm constitutes a challenge for training ML models in pharma as the solution to certain process problems requires different algorithm or combination to achieve an optimal goal (Wuest *et al.*, 2016). Although, efforts have been made to define "general ML techniques," the variety of issues and their demands emphasise the necessity for specialized algorithms with advantages and disadvantages.

Another challenge of ML encountered in pharma is data acquisition. This is also a limitation as the availability, quality, and composition of the pharmaceutical manufacturing data hand have a strong influence on the performance of ML algorithms (Pham and Afify, 2005; Wuest *et al.*, 2016). The dataset may present some challenges, such as high-dimensional data, which can be problematic for certain machine learning algorithms due to its high level of redundant and irrelevant information that can affect how well learning algorithms work. Hybrid ML techniques (a combination of several algorithms) also contributes to the challenges encountered in pharmaceutical process optimization by applying multiple algorithms to yield better outcome rather than implementing a single algorithm (Lee and Ha, 2009). This is becoming prevalent and is expected to further increase the complexity currently present. Interpreting the findings presents another challenge as it must be considered that the parameters of the selected algorithm, the parameter settings, the planned outcome, the data, including its pre-processing, and the format or illustration of the output are all significant to the intended outcome (Quadrianto and Buntine, 2017; Widodo and Yang, 2007).

2.6. Research Gap

In preparation of smart manufacturing, the real-time applicability of ML in pharmaceutical manufacturing processes needs to be studied, which, to my best knowledge, has not been extensively detailed in previous review papers yet. Most publications relating to use of machine learning in the pharmaceutical industry is actively involved drug research, discovery and model designing leaving model utilization, and accessibility overlooked. This research paper will aim to fill in the gap by conducting a survey on the current and future uses of ML in different pharmaceutical manufacturing areas. In addition, the aspect of machine learning integration with existing infrastructure and workforce would also be explored to maximize the potential for widespread adoption and impact. Addressing these gaps would not only advance the use of ML in pharma but also facilitate the development of practical and effective solutions that can enhance efficiency and quality across the pharmaceutical industry.

2.7. Literature Conclusion.

Based on the examination of journal articles and books published between 2010-2023, the use of ML has gained widespread application in the pharmaceutical and biopharmaceutical manufacturing sector and is likely to continue at a rapid pace. Notably, the assessments concerning the benefits pursued and achieved by applying machine learning in the pharmaceutical manufacturing sector has received minimal attention in semi-solid, liquid, and parenteral manufacturing in the aforementioned reviews. Therefore, research evaluating the maturity of the proposed solutions and emphasising the specific and tangible advantages of implementing ML approaches in these areas is needed. According to most publications, supervised-based algorithms are superior in most manufacturing applications. However, with the rapid growth in available data due to improved sensor technologies and more awareness, unsupervised learning would increase in significance in pharma. Further research is needed to focus on the developing machine learning techniques tailored to the specific requirements and constraints of pharmaceutical manufacturing, enabling manufacturers to understand and validate the recommendations provided by the model.

CHAPTER THREE

3. Research Methodology

3.1. Overview

Research has been defined by various authors in different ways. It always begins with a question, or a problem and its purpose is to find answers to questions through the application of systematic and scientific methods. Thus, research methodology is the systematic approach towards purposeful investigation. This needs formulating objectives, collection of data on relevant variables, analysing and interpreting the results and reaching conclusions whether in form of a solution or certain generalizations (Bhattacharyya, 2006; Davidavičienė, 2018). The pharmaceutical industry plays a critical role in the global healthcare sector by developing and manufacturing live-saving medications and treatment. In recent years, there has been a growing interest in leveraging machine learning techniques to address the challenges often faced in the industry and optimize pharmaceutical manufacturing processes. By harnessing the power of ML algorithms, pharmaceutical companies can streamline production, improve product quality, reduce waste, and enhance overall efficiency. Given the increasing pressure to accelerate drug development and improve manufacturing efficiency, there is a pressing need for research that explores the application of machine learning techniques. Traditionally, pharmaceutical manufacturing processes have relied on manual interventions, empirical adjustments, and predefined control strategies to optimize production efficiency and maintain product quality. While these methods have been effective to some extent, they lack the agility, adaptability and predictive capabilities required to address the dynamic nature of modern manufacturing environments.

Chapter 1 and 2 has provided the background, relevant literatures on several uses and application of machine learning algorithms for process optimization in specific areas of pharmaceutical manufacturing. This chapter provides an overview of the research methodology applied in this thesis as well as to describe the approach taken to answering the research questions that relate to the established study objectives to ensure the reliability and validity of this research. The aim of this study is to identify key areas of manufacturing that ML could potentially improve in the pharmaceutical sector by evaluating the knowledge and awareness of operators and production managers regarding the processes involved in manufacturing in the pharmaceutical facilities.

To fulfil the study objectives, a suitable research technique was developed utilising the Research Onion as shown in figure 3.1, a widely acknowledged research tool that gives the basis of a theoretical concept which involves determining an acceptable method, philosophical perspective, and approach to the study. The research onion provides a rather exhausting description of the main layers or stages which are to be accomplished to formulate effective research methodology (Melnikovas, 2018).

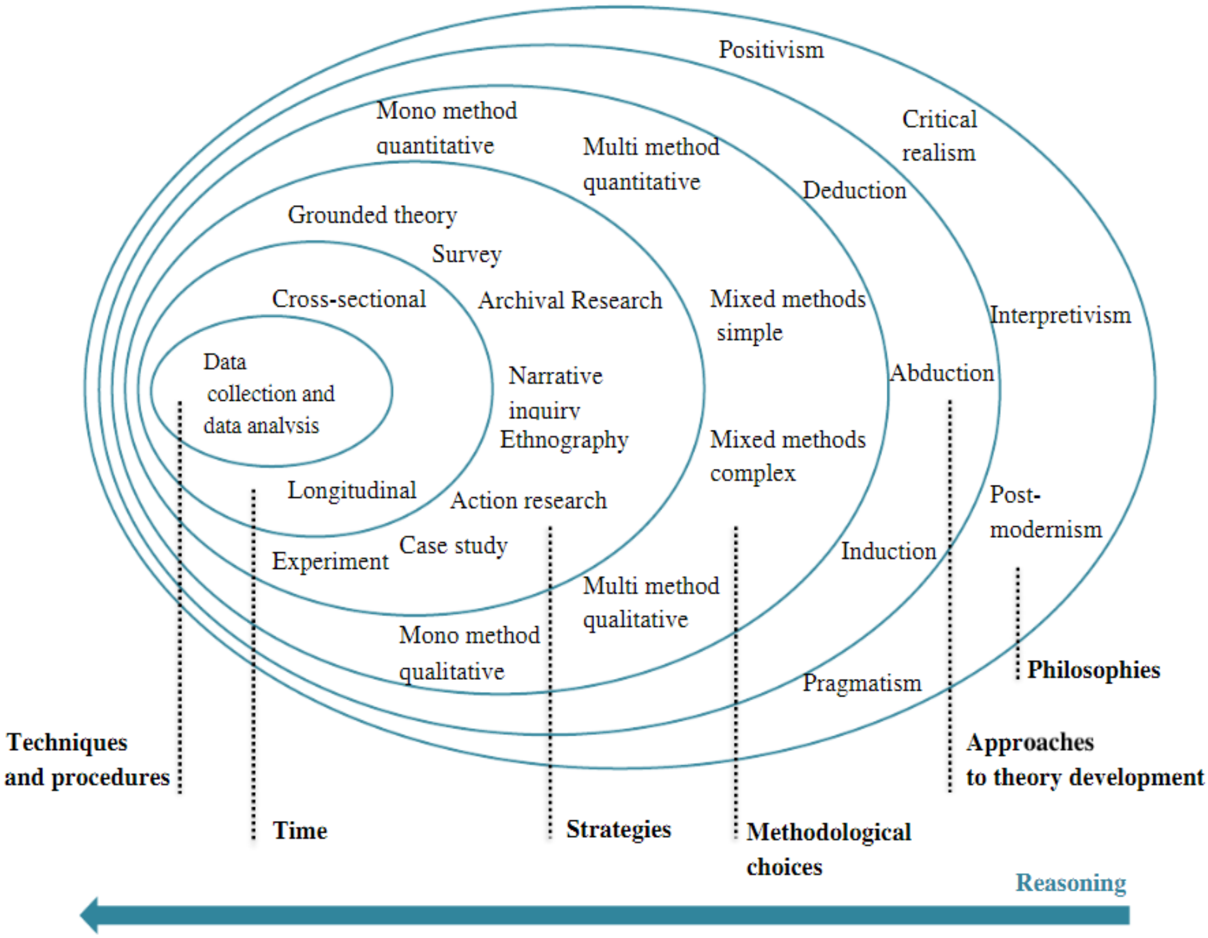


Figure 3.1 The research onion framework (Saunders et al., 2019).

3.2 Research Process

According to Alphonse (2023), different authors proceed in a variety of ways when it comes to conducting research and they opt for a research process depending on the nature of research, the type of data to collect, approaches adopted for data analysis, or techniques employed to come up with research conclusion. The approach used in this study is based on the instrument used for data for data collection and analysis and how it provides significant insight for the primary research. The procedures followed to complete this research is depicted in figure 3.2 and the framework is based on Safen and Gustavsson's (2020) research procedure. Starting with the concept of artificial intelligence and machine learning techniques in the identified pharmaceutical manufacturing research areas to the formulation of the research questions. A systematic literature review analyses the state of art pertinent to the formulated research questions. Based on this existing knowledge, the study design is evaluated.

Time horizon- As stated by (Saunders et al., 2019), a typical research timeline can either be;

- Cross-sectional study – Data collected from a specific group at a single point in time.
- Longitudinal study - Data collected from the same group repeatedly overtime.

Due to the time constraints of primary data collection for this research, the study would employ a **cross-sectional time horizon.**

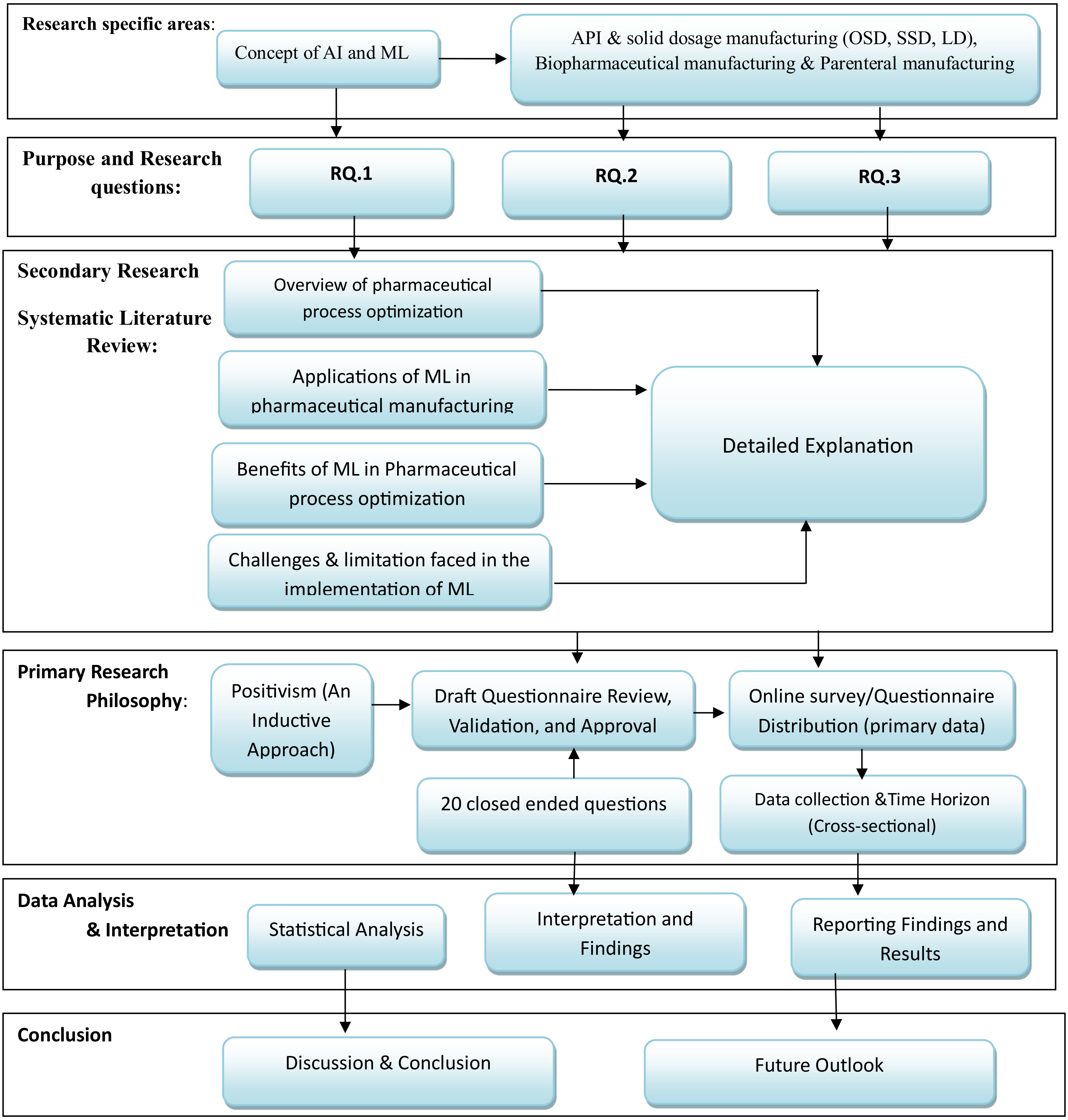


Figure 3.2 Diagram illustrating Research Process Flow Chart

3.3 Research Philosophy

The conceptual framework on assessing the specific factors and/or emerging technologies that contribute to the integration of machine learning algorithms in pharmaceutical manufacturing processes is based on a philosophical approach and the use of the right philosophy hold much significance in the execution of this research. These philosophies are typically studied from three perspectives: Ontology, Epistemology or Axiology. As proposed by Melnikovas (2018), Ontology is the assumption about the reality faced in the research and how existence is understood, Epistemology combines positivism, realism and interpretivism to determine the validity of research data and its acquisition methods and Axiology focuses on the impact of viewpoint sand principles on research data gathering and evaluation. It explains how public opinion influences a study's perspective and variables (Tengli, 2020).

The philosophical approach to this primary research strategy would entail a critical analysis of the nature of knowledge formation, ethical implications, and the basic reality of applying machine learning into the manufacturing process. This study will adopt positivism where survey-questionnaire will be used to collect quantitative data to provide quantifiable and objective insights regarding the use and effects of machine learning in dosage manufacturing, biopharmaceutical, and parenteral manufacturing processes. Positivism was employed to generate a suitable conclusion. It was implemented by acquiring quantifiable observation that aided statistical interpretation or analysis of data acquired from willing participants (Saunders et al., 2019). It will be used to comprehend the participants subjective experiences, opinions, perceptions, and interpretation during quantitative data collection with process operators and production managers, in selected pharmaceutical manufacturing facilities. As mentioned in the research objectives, the study would investigate the impacts of crucial processing factors that are essential to the manufacturing plant and may also examine the current challenges influencing the process and product lifecycle.

- To determine the material attributes and quality parameters that majorly impacts the critical quality attributes of the final product in real time.
- And to determine the areas of improvement using machine learning to streamline operations, the advantages it offers and its impact on decision making.

3.4. Research Design

As described by Asenahabi (2019) and Cresswell and Cresswell (2014), an effective method for interpreting the data analysis is a crucial component of a research design, as it ensures sufficient findings and correlations from the study. A research design is a plan, structure and strategy of investigation that is adopted with the aim of obtaining answers to research questions with optimal control of variables (Saunders et al., 2019). As shown in figure 3.3, a research design is divided into three groups: quantitative, qualitative, and mixed method approaches. This study aims to address three primary areas: the Concept of Machine Learning, Process Optimization, and Pharmaceutical Manufacturing and how they relate to a suggested conceptual idea. The Literature and expert opinions collected shed light on these areas and explains what must be done in terms of the research design to accomplish the study goals. The study will use an inductive philosophical method to develop a conclusive theory based on the systematic literatures reviewed and the research methodology.

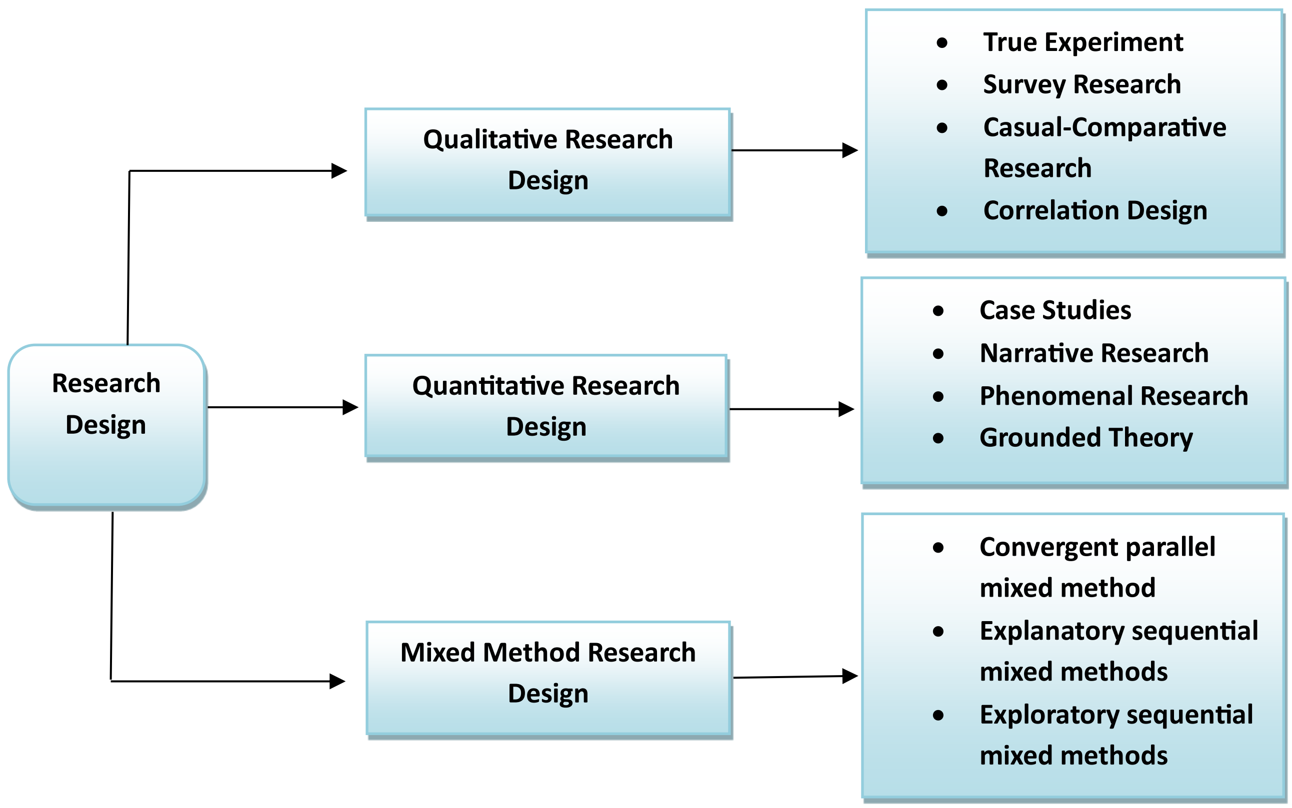


Figure 3.3 Diagram Illustrating the Research Process (Asenahabi, 2019)

3.5. Research Strategy

The strategy for gathering primary data will involve an exploratory methodological approach which would include a design to gather quantitative data. To proceed with gathering relevant knowledge for primary research in the above mentioned pharmaceutical manufacturing areas, an assessment to gain full insight and information on the current challenges in pharmaceutical manufacturing areas, explore specific applications and identify opportunities for optimization through machine learning will be done and this can be accomplished by using a standardized online survey-questionnaire with closed-ended, clear and concise questions to collect data on the current use of machine learning in the above mentioned pharmaceutical manufacturing areas. This was chosen because it facilitates the evaluation of the perspectives of manufacturing personnels actively involved in pharmaceutical manufacturing processes and they include process operators, production managers, quality control specialist, industry expert and research specialist, from many participants.

As there are three (3) research questions guiding this research, the survey will be conducted with clear research objectives and precise questions. For the first and third question, the study intends to find out know the rate at which the application of machine learning enhance operations of pharmaceutical manufacturing and the specific factors or emerging technologies that contribute to this AI-driven optimization approach and how do they differ across different pharmaceutical facilities and to determine if using machine learning used in these manufacturing areas can assure public safety together with the ethical and regulatory considerations. To answer these questions, production managers and regulatory experts specializing in dosage formulations would be chosen to gain understanding of their firsthand experience, perspectives, and insights on the formulation processes.

The second research question seeks to know the opportunities associated with the adoption of machine learning in the pharmaceutical sector, specifically in the aspect of quality attributes, uniformity, yield, and compliance with regulatory standard. To answer this question, a survey-questionnaire (multiple-choice questions) focusing on the prevalence of ML adoption and the types of algorithms used would be distributed to the participants in selected pharmaceutical manufacturing facilities. This qualitative approach would identify the trends and patterns of potential or already existing benefits of machine learning.

A timetable for progress milestones was designed for effective time management during the period of the research using a Gantt Chart

	JANUARY	**FEBRUARY**	**MARCH**	**APRIL**	**MAY**
Refining research objectives					
Literature					
Methodology					
Research design & Ethics Approval					
Collection of primary data					
Statistical Analysis and Interpretation of data					
Discussion					
Final research write up					

3.6. Questionnaire Design.

With surveys, gathering data from a large number of participants is made straightforward and standardised, which led to designing a questionnaire for this research. The structured questionnaire was created using Google form on one-drive consisting of 20 closed-ended question for quantitative data. Open ended comments were included at the bottom of some questions to allow for participant additional knowledge, and this was conducted during spring of 2024, by sending an online invitation (survey's URL) to selected participants via professional platforms such as LinkedIn profile, emails, and social network such as Facebook, and WhatsApp. The questionnaire is divided into three (3) sections. The first section which seeks to obtain demographic information and background information on their knowledge of different manufacturing procedure using machine learning and other enabling technologies employed in their manufacturing plant.

The second section of the questionnaire goes deep into the context to gain a better understanding of the participants knowledge of the type of algorithm used to control process and product variability during the manufacturing process.

The third section seeks to investigate the critical process parameters that are integral to the manufacturing plant and to determine the areas of improvement using machine learning to streamline operations. The survey will use different question styles to gather data, including dichotomous questions, multiple choice questions and the Likert-type interval scale questions. The questionnaire was designed with the use different question styles to gather data, including dichotomous questions, multiple choice questions and the Likert-type interval scale questions. This was written in English and answering the survey took about 10minutes.

3.7. Research Population and Sampling Technique

Within this study, choosing a sample size is required in order to answer the established research questions. The first stage in the sampling process is to clearly define the target population and a sampling technique to ensure the validity and reliability of the results. The sample size is based on the number of responses and not necessarily the number of questionnaires distributed (Taherdoost, 2016). Considering the diverse encompassing stages of pharmaceutical manufacturing, a stratified sampling technique was deemed appropriate where the population was divided into different strata based on factors such as different process stages. As a general guideline, a minimum sample size of 70-90 participants was taken into consideration. This figure is adequate to provide the degree of precision and clarity for the data analysis. The following categories of participants were selected and given access to survey.

- Process Operators and Engineers: they play a major role in designing, optimizing, and troubleshooting manufacturing processes in the pharmaceutical industry. They have in-depth knowledge of process parameters, equipment, and production techniques.
- Production Managers: Manufacturing managers oversee the day-to-day operations of pharmaceutical manufacturing facilities. They have a comprehensive understanding of production schedules, resource allocation, and operational challenges. Their perspectives on use of machine learning in manufacturing processes are essential for evaluating feasibility and scalability.

- Industry Experts: their collaborative effort with artificial intelligence maturity model aims to drive initiative relating to machine learning, process optimization and regulatory compliance in pharmaceutical manufacturing.
- Quality Assurance/Control Specialist: Quality assurance specialists are responsible for ensuring compliance with regulatory standards, maintaining product quality, and implementing quality control measures throughout the manufacturing process. Their insights into quality management practices are valuable for assessing the impact of machine learning on product quality and regulatory compliance.

3.8. Inclusion and Exclusion Criteria

The requirement for study participation was proficiency in English, prior knowledge and experience in pharmaceutical manufacturing, there was no regional specification or geographical restriction for individuals who participated in the survey. Vulnerable participants and individuals below the age of 18 was excluded from the study as well as individuals who lost interest in completing the survey.

3.9. Data Analysis Strategy.

This part of the research process is very critical as when it is not very well performed, it may provide misleading results. Data analysis involves identifying and comparing the outcomes of various treatment upon several groups, understanding their underlying causes or influences **of** a research problem and making a decision to achieve research goals (Kothari, 2004; Dibekulu, 2020). As my research aims to review various aspects of pharmaceutical manufacturing operations that machine learning algorithms can potentially improve, my strategy for analysing each category of collected primary data to ensure accuracy, completeness is as follows:

- The quantitative analysis and design space would rely on regression method (inferential statistics) to show the frequency and probability of a variable occurring.
- Descriptive statistics would be employed to summarize the key findings.
- Cross tabulations would be explored to understand the relationship between variables such as the correlation between the use and these technologies for the data collected from different facilities.

- Analytical software would be used for data visualization to design tables, charts, and graphs for analysis to identify patterns, gaps, and pinpoint areas for further research.

By employing these techniques for data analysis, the collected survey data for the study would be examined, leading to the acquisition of useful insights that may enhance the comprehension of the utilization of machine learning in pharmaceutical manufacturing processes. The primary data for this research was gathered at once due to its cross-sectional time horizon.

3.10. Ethical Considerations

A research study must ensure the protection of human participant, by adhering to relevant ethical guidelines (Arifin, 2018). Due to the in-depth nature of the study process, ethical considerations are especially relevant in quantitative research. Prior to data collection, the strategy to data sourcing was ensured, in compliance with the general data protection legislation (GDPR) and the ethical guidelines laid down by Griffith College's ethics and research committee. In addition, efforts were made to ensure that all participants involved fully understood the purpose of data collection, how their data would be used, the confidentiality of the information provided and how it will be utilised solely for the purpose of this dissertation. This was done by providing a comprehensive description of the research purpose and obtaining an informed consent from each participant to ensure voluntary participation (this was provided in English Language). Prospective participants were made fully aware of their right to withdraw from the study at any time, even after they had signed the informed consent.

Additionally, no personal information was collected therefore it was impossible to identify the respondents to prevent bias during data collection and analysis and the survey data was not manipulated or falsified.

Furthermore, the anonymity and confidentiality of the participants involved in this research was preserved by not revealing the identities during the data collection and analysis process. The primary research data was collected and stored in a structured manner for convenient retrieval and analysis on my personal computer (secured with a strong password), and on one drive for backup purposes and this will be retained for a period of 2 years until deemed appropriate for disposal by the authority of Griffith College.

3.11. Conclusion

This chapter has covered the general differences between the various stages of research methodology in detail. It is now necessary to indicate that further investigation or in-depth analysis of the approaches is required. This can be done by highlighting the casual relationship and discussing relevant points in the analysis of the data collected that will be further examined in the following chapter in order to arrive at the overall research conclusion in the final chapter.

CHAPTER FOUR

4. Data Analysis and Interpretation

4.1. Introduction.

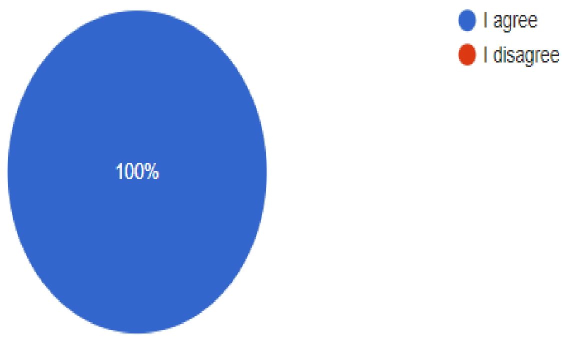
This chapter provides an analysis and interpretation of the primary data obtained via survey that was administered through a direct link for confidentiality between the 28th of March - 22nd of April 2024 (a total of 26 days) relating to the use of machine learning algorithms for process optimization in pharmaceutical manufacturing industry. The sample size included production managers, industry experts and manufacturing professionals with relevant experience in product development and regulatory affairs in pharmaceutical companies.

4.2 Analysis of Responses.

The beginning of the survey provided a brief description to allow the participants to perceive the content of the questionnaire and was divided in three sections as well as an agreement message (consent) for willing participants to proceed as seen in Table 4.1. All participant's responses were recorded in a spreadsheet and was evaluated with Descriptive statistics using Microsoft Excel. A total of 69 responses was gathered providing an overall response rate of 76.6%. The average response rate was determined using central tendency measures, and any possible association between variables was examined using cross tabulation.

Table 4.1. Participant Consent

CONSENT	RESPONSES	PERCENTAGE
AGREE	69	100
DISAGREE	0	0
TOTAL	69	100



The majority of participant's responses came from 27 manufacturing personnels (a total of 39.1%) that was sub-divided into process operators/engineers, process technicians and production managers, followed by 20 quality assurance/control specialist (a total of 27.5%), 15 industry experts which was further divided intro data scientist, R&D professionals (yielding a total of 21.7%), and lastly 8 research specialist (yielding a total of 11.6%). All participants willingness to engage in the survey is indicated by the 100%.

4.3. Quantitative Analysis

Section 1 (Background and Demographics Data on Survey Respondents) - To better assess the participants' level of knowledge, this section collected data on their role of profession, and their years of experience in selected field.

Q1- Which of the following best describes your current role of profession?

The distribution of respondents in different roles in pharmaceutical manufacturing industry highlights the widespread interest and potential application of ML for process optimization within different facilities.

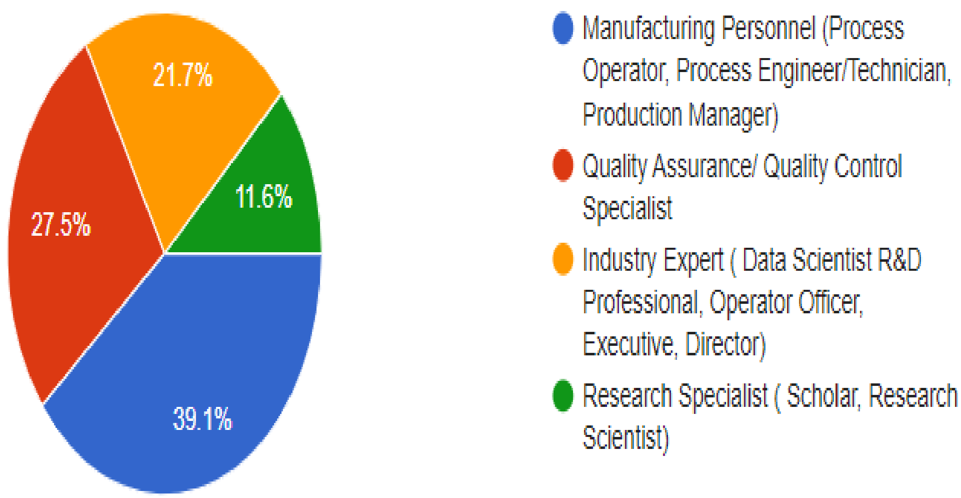


Figure 4.1 Analysis of survey demographics according to profession.

Table 4.2. Total Number of Responses by Profession

Profession	Responses
Manufacturing personnel	27
Quality Assurance / Control	19
Industry Experts	15
Research Specialist	8
Total	**69**

With manufacturing personnels having the overall highest response rate (39.1%), figure 4.1 and table 4.2 reflects the cross-functional nature of leveraging ML for process optimization as the individuals in this selected group are directly and actively involved in product formulation and development from raw material selection, product, and process lifecycle to finished product.

Q2 – Please choose the option that most accurately reflects your years of experience in the selected given role.

Respondents were asked to provide their years of experience to determine the duration of their practice in selected roles that could provide relevant insight to fulfil research objectives. The selected years of experience was between less than a year – above 10 years as shown in figure 4.2.

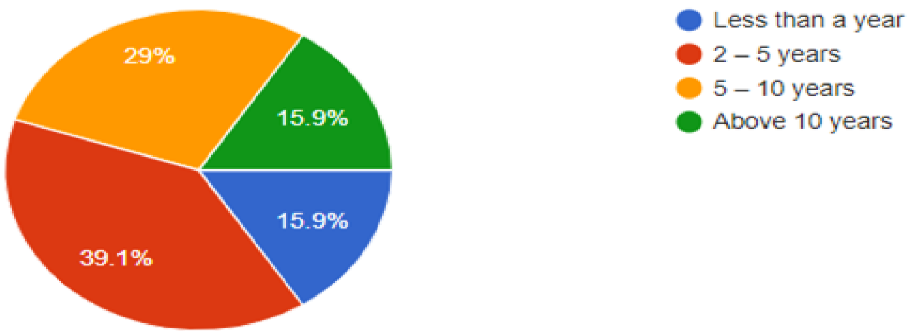


Figure 4.2. Analysis of Participants' years of Work Experience

Table 4.3 Overall Response Rate of Participants Work Experience.

Selected Profession	**Years of Work Experience**				**Respondents**	**Response rate (%)**
	< 1yr	**2-5yrs**	**5-10yrs**	**>10yrs**		
Manufacturing personnel (process operators, engineers	4	9	9	4	26	37.6
Quality Assurance/Control Specialist	1	11	6	2	20	28.9
Industry Experts (Data Scientist, R&D)	2	4	5	5	16	23.2
Research Specialist (Scholars, Research Scientist)	3	2	1	1	7	10.5
Total Survey Responses	10	26	21	12	**69**	**76.6**

As shown in figure 4.2, and table 4.3 majority of the respondent are within 2-5 years of work experience with a total of 39.1% in pharmaceutical manufacturing. A significant number of participants also possess experience beyond 5 years (29%) and a small percentage of 15.9% respondents possess experience above 10 years and less than a year. Since most of the extensive experience lies between 2-5years and 5-10 years, relevant insights can still be obtained from this analysis. Furthermore, the tabular data generated a total of 69 respondents that completed the questionnaire with an overall response rate of 76.6%.

Q3. How familiar are you with artificial intelligence (AI) and machine learning (ML) technologies?

This question is focused on assessing the familiarity of participant's opinions on the types of ML algorithms used generally as the use AI and ML have gained global population over the years specifically in the manufacturing sector. Options on participant's level of familiarity - extremely familiar, somewhat familiar, or completely unfamiliar were provided to them.

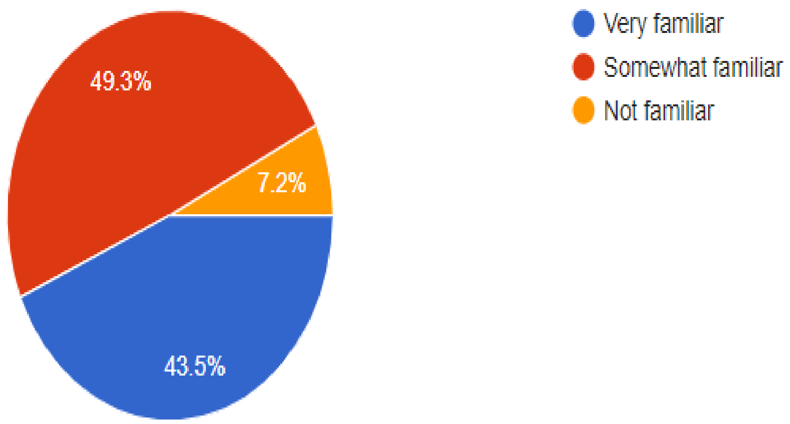


Figure 4.3. Pie chart on Participant familiarity of ML techniques.

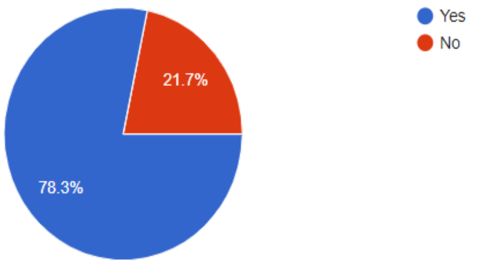
According to the analysis as shown in figure 4.3, 34 of the participants (quality assurance/control) responded to being somewhat familiar with ML technologies yielding a total of 49.3%, 30 participants (manufacturing personnel, industry experts and research specialists) responded to being very familiar (43.5%) and interestingly, 5 participants were completely unfamiliar yielding a total of 7.2%. However, it was discovered that the work experience of majority of the respondents that were found to be very familiar and somewhat familiar lies between the 5-6 years and 2-5 years respectively. This implies that the selected participants are aware of machine learning technologies and therefore are a great audience to comment on the use of ML in pharmaceutical manufacturing processes.

Q4. Have you heard about the application of ML algorithms in pharmaceutical manufacturing processes prior to this survey?

To better understand the familiarity of participants on the use of ML for process optimization, this question seeks to inquire about the participant's connection between the use of ML in a specific process and how they view the application in their field of professional development by providing a YES or NO option.

Table 4.4 Familiarity with AI and ML technologies.

Responses	Frequency	Percentage
YES	54	78.3%
NO	15	21.7%
Total	69	100%



With descriptive statistics, a significant number of 78.3% responded YES to being aware prior to the survey therefore their opinions would offer insights to this research. This implies that majority of respondents has a positive view on using ML within their manufacturing facility. On the other hand, 21.7% responded NO to this question which only gives the suggestion that these individuals may not have encountered the concept of ML, or they are currently implementing other types of technologies in their organization.

Q5. To what extent do you believe that improving process optimization in pharmaceutical manufacturing may be accomplished using machine learning?

To answer the first research question, Q5 & Q6 question was asked to see if respondents believe that using machine learning is beneficial to their organization to maximize efficiency, quality, and yield. Their responses would determine the level of ML adoption in their organization. See fig. 4.4 and fig. 4.5.

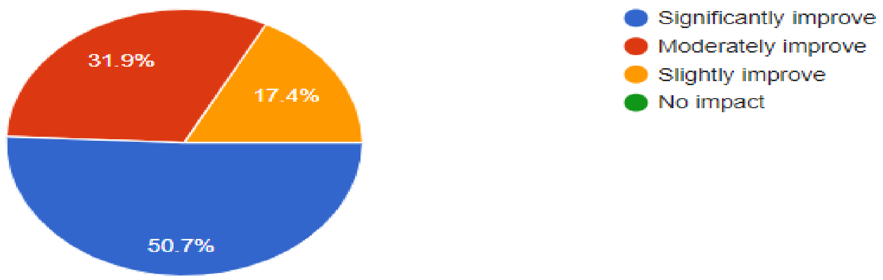


Figure 4.4 Analysis of Participant's Perception on the use of ML

50.7%, (35) a high percentage of obtained responses are of the belief that machine learning would significantly improve manufacturing processes if applied during process optimization. Majority of the individuals with this belief are within the category of 5-10 years' work experience. Results from this analysis also shows that certain respondents with a total of 31.9% (22) also believed that ML can moderately improve manufacturing processes and 17.4% (12) were a bit sceptical towards it by picking the 'Slightly improve' option.

Q6. Which of the following do you believe best describes the current adoption of machine learning technologies in pharmaceutical industries? This statement was asked to follow up on Q5 to further assess participants knowledge on the current application of machine learning.

Figure 4.5 Analysis of Participant's Perception on the use of ML

Interestingly, majority of the participants believe that machine learning can significantly improve manufacturing processes in the industry as shown in fig.4.4. However, only 21 (30.9%) are of the impression that ML algorithms is strongly being implemented in the pharmaceutical industry. A response rate of (63.2%) was generated from 43 of the participants with the impression of ML being moderately implemented and 5.9% from 4 participants that believes that ML is not implemented in the industry. The basis of these responses simply reflects on the current use of ML and how it the differs across organizational context.

4.4. Quantitative Analysis on Participant's Awareness of ML in pharmaceutical manufacturing processes

The second section of the survey is directed to the second research question "What are the opportunities associated with the integration of machine learning in the pharmaceutical sector, specifically in the aspect of quality attributes, uniformity, yield, and compliance with regulatory standards?" of this thesis by evaluating each participants' knowledge and data obtained about their familiarity with the concept, whether they have heard of ML prior to the survey and whether they are actively implementing any ML technique in their facility was collected.

Q7-Q9 What do you believe are the primary motivations for implementing machine learning technologies in pharmaceutical manufacturing and state the type of algorithm used if there is any you have implemented to optimize a process.

The expectations from this part of the questionnaire are to tailor the responses based on analysed literature by providing answers in checkboxes where respondents could choose more than one option as they see fit. see in figure 4.6.

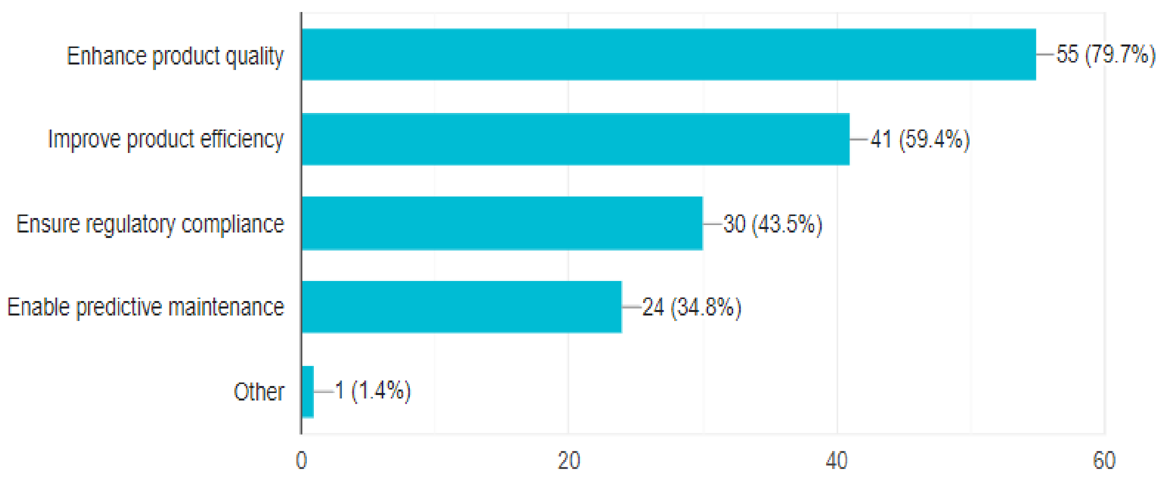


Figure 4.6 Participant's opinion on the benefits of applying ML.

As stated in most literature reviewed in the study, the overall benefits of ML in pharmaceutical manufacturing is to achieve process efficiency and product quality and majority of the participants would agree with the phenomenon as 79.7% & 41% accounted for a total of 55 and 41 counts respectively. Additionally, 43.5% accounted for 30 counts on regulatory compliance and 34.8% accounted for 24 counts for predictive maintenance.

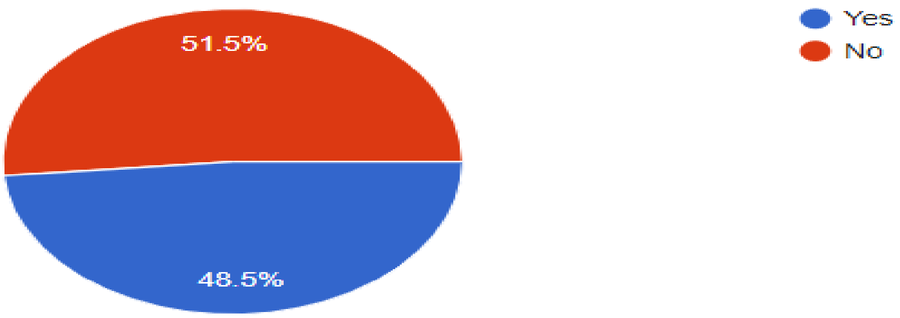


Figure 4.7 Response on the use of ML-based optimization technique.

According to the analysis above, 51.5% responded 'NO' indicating that a significant portion of the participants has not implemented any machine learning-based approach to optimise processes during manufacturing. The participants that responded "YES" has a response rate of 48.5% and was asked to specify on the type of algorithm used as shown in figure 4.8.

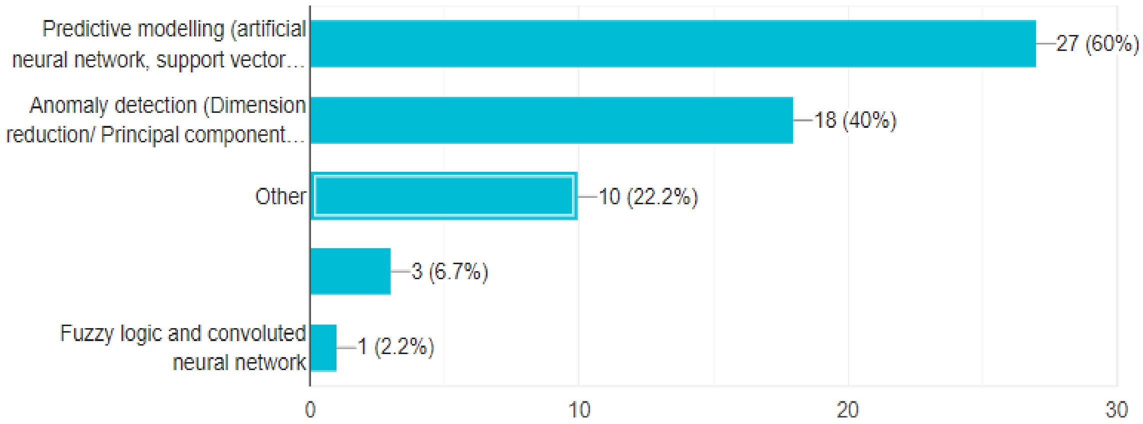
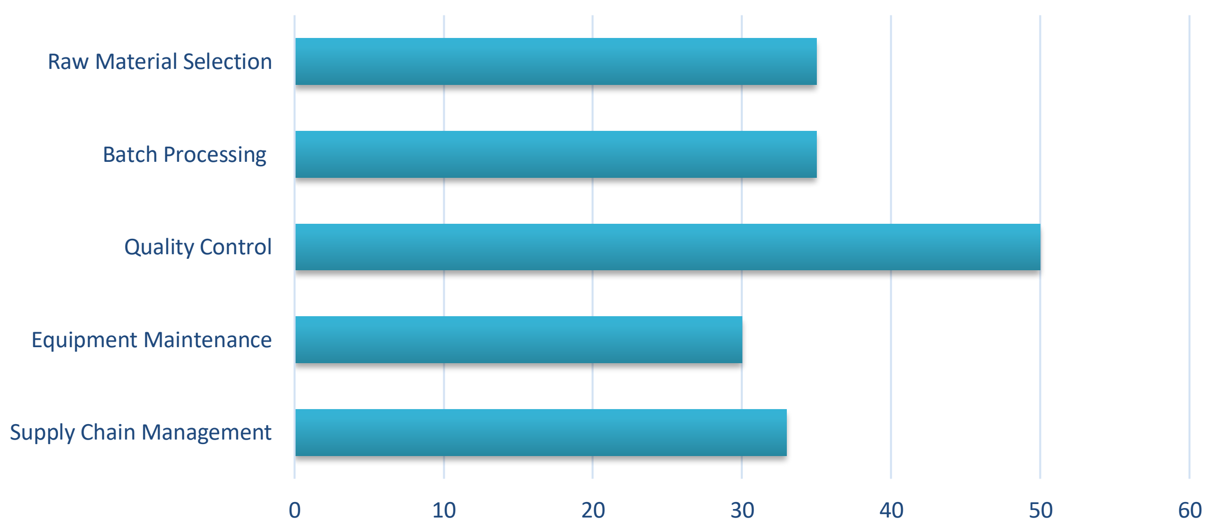


Figure 4.8 Analysis of Participant choice of selected algorithm.

This was a combination of a closed and open-ended question where respondents could specify the type of algorithm used that was not stated in the option. A total of 27 counts (60%) confirmed using predictive analytics in their organization. This option had a variety of algorithms such as ANNs, SVMs, decision trees and random forest provided suggesting that majority of the utilised algorithms used are supervised-based. A total of 18 counts (40%) confirmed using certain unsupervised based techniques for anomaly detection such as clustering analysis, PCA and Dimension reduction. In an open statement, 1 respondent mentioned Fuzzy logic and convoluted neural network.

Q10 Which aspect of pharmaceutical manufacturing process do you believe can benefit the most from machine learning-based optimization?



	Supply Chain Management	Equipment Maintenance	Quality Control	Batch Processing	Raw Material Selection
Response Count	33	30	50	35	35

Figure 4.9. Specific areas ML could potentially improve according to participants.

In assessing certain areas that tends to benefit from the opportunities that ML has to offer, It was observed from the Bar chart, a high rate of 73.5%(50 counts) from respondents mostly in the field of industry experts and quality assurance specialist that were in favour of quality control, followed by raw material selection and batch processing with a rate of 51.5 % each (35 counts), a rate of 48.5% (33 counts) in favour of supply chain management and 44.1% (30 counts) in favour of equipment maintenance.

We have learnt from literature that ML uses pharmaceutical data generated by PAT to analyse processes in real time which begs the question regarding the availability of data in different facilities.

Q11. How would you rate the availability and quality of data required for machine learning based optimization based on your experience?

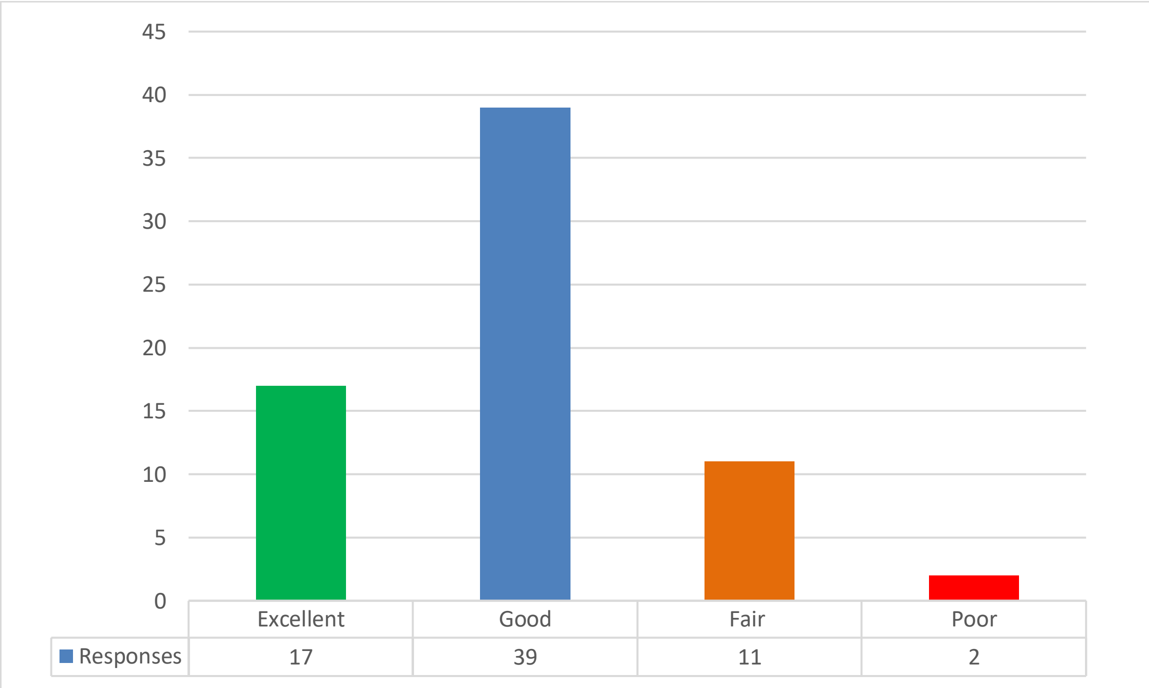


Figure 4.10. Participants' knowledge on Data Quality and Availability

A good number of respondents, 56.5% rated data quality "Good", 24.6% of respondents rated data quality "Excellent", 15.95% of respondents rated "Fair" and 2.9% rated "Poor". High quality data is of major importance for the incorporation of machine learning and majority of responses from this question implies in the availability of data within facilities.

Consideration was given to public safety regarding the use of ML in pharmaceutical manufacturing. Despite the benefits of ML to streamline formulation development process, it must be determined whether there are risks associated with its use. The next question is categorised by the level of agreement ranging from "Strongly agree to Strongly disagree".

Q12. Do you believe that the use of ML in pharma can assure public safety?

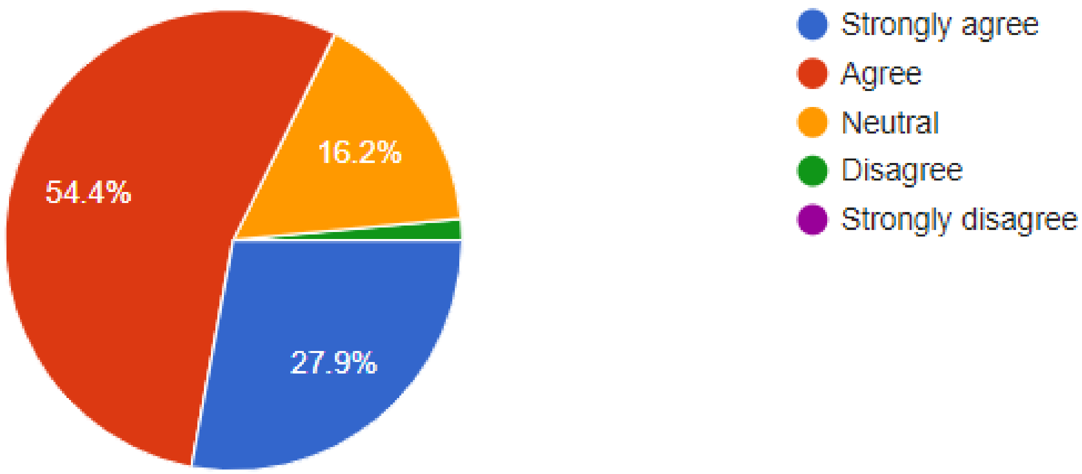
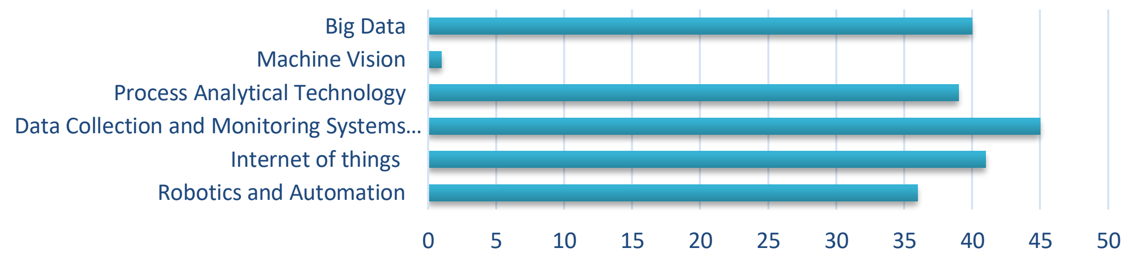


Figure 4.11. Paricipant's viewpoint on machine learning for public safety

The findings from this analysis is categorised by the persentage rate of responses in each radial slices of the piechart. The results showed that 54.6% (37) ageed to the statement where manufacturing personnels and quality assurance specilists shared the same opinion, 27.9%(19) were in strong agreement, 16.2%(11) were neutral and 1.5%(1) did not rely on the statement and completely disagreed. With majority of the reponses in the "agreed" and "strongly" agreed category, this suggests that ML in pharma has a positive impact on public health.

Q13. Which enabling technology do you believe contributes to the use of ML?

The first research question also aimed to determine the technologies that facilitates machine learning implementation in the pharmaceutical industry. Answers to this question listed other Industry 4.0 enabling technolgies that contributes ML-based approach and was provided in checkboxes where paticipants could selected more than one option. See Fig. 4.12.



	Robotics and Automation	Internet of things	Data Collection and Monitoring Systems (Sensors and Probes)	Process Analytical Technology	Machine Vision	Big Data
Response Count	36	41	45	39	1	40

Figure 4.12. Respondents' choice of technologies that contributes to ML-Bases approach.

Several studies cited in literature have indicated that real-time data is generated through PAT and IoT enables the collection of data from smart devices and systems that machine learning analyses. With a high response rate of 66.2% (45), 60.3% (41), 58.8% (40), 57.4% (39), and 1.5% (1), the spread of selected technologies among participants is in correlation with reviewed literature with IoT having the highest range and machine vision having the lowest range.

Q14. How would you rate the ability of machine learning to reduce cycle time, variability, and resource consumption in manufacturing processes?

This question was provided on a scale of 1-5 with 1 indicating less likely, 3 indicating neutral, and 5 indicating more likely on a 5 scale Likert rating format.

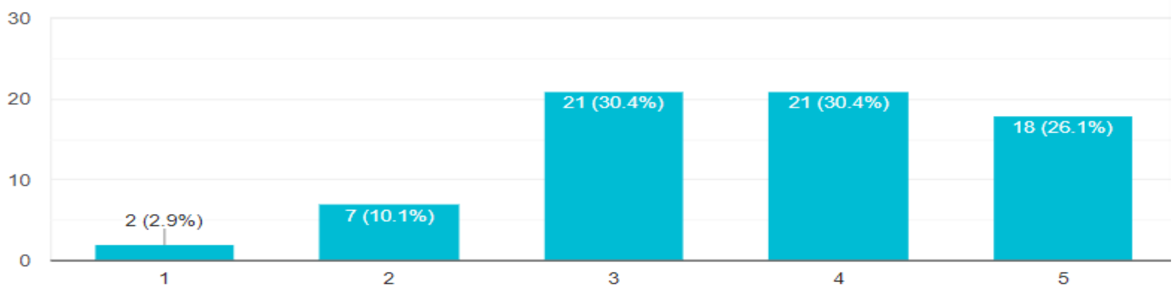


Figure 4.13. Respondents' Ranking on ML Abilities

How would you rate the ability of machine learning to enhance drug release profile prediction?

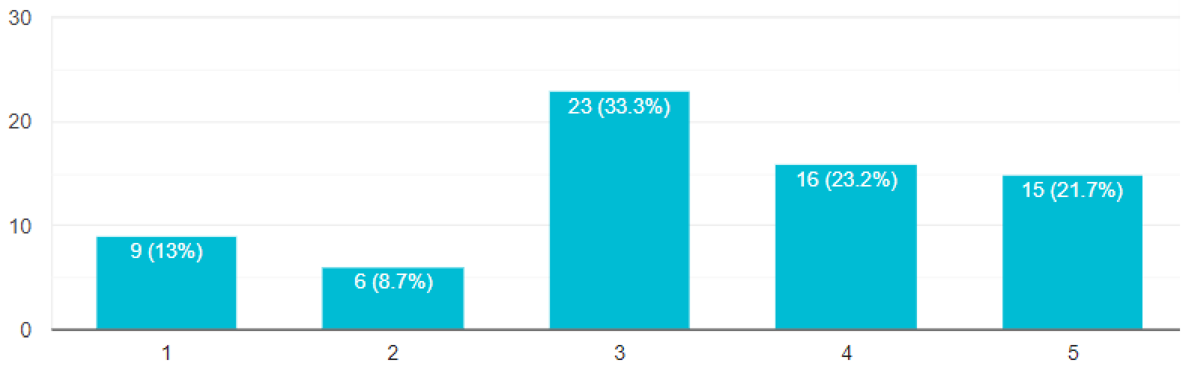


Figure 4.14. Respondents' Ranking on ML Abilities

How would you rate the ability of machine learning to optimize drug formulation, stability, bioavailability, and efficacy in manufacturing processes?

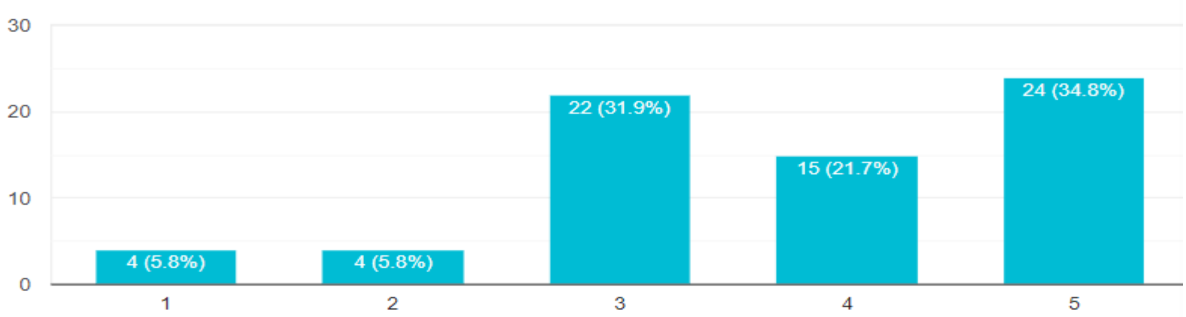


Figure 4.15. Respondents' Ranking on ML Abilities.

The Major challenge commonly faced in pharmaceutical manufacturing is variability. On a micro-level, the complexity of precise drug formulation and its interaction with excipients and APIs to achieve the desired product and on a macro level by considering parameters from raw materials to finished products given the variables involved (Lou *et al.*, 2021). In an attempt to evaluate participant's perception on how ML can help alleviate this challenge, the above questions were asked, and in the analysis, it was observed that;

- 26.1% (18) were confident in the abilities of ML and rated more likely, 30.4% (21) were mildly confident, 30.4% (21) were neutral towards the statement and 2.9% (2) rated less likely in the first chart.

- 21.7% (15) rated more likely, 23.2% (16) were mildly confident, 33.3% (23) were neutral and 13% (9) rated less likely in the second chart.
- 34.8% (24) were very confident and rated more likely, 21.7% (15), were mildly confident, 31.9% (22) rated neutral and 5.8% (4) rated less likely in the third chart.

4.5 Quantitative Analysis on CPPs in specific manufacturing areas

Here, information was gathered on the participant's familiarity with CPPs that are integral to the manufacturing plant to explore the current challenges influencing process and product lifestyle and how machine learning can improve the processes.

Q.15. How familiar are you with the concept of critical control parameters in pharmaceutical manufacturing?

To identify the critical material attributes and quality parameters that significantly influence the critical quality attributes of the final product in real time. This question aims to assess the validity of the statement by directing it to the participants.

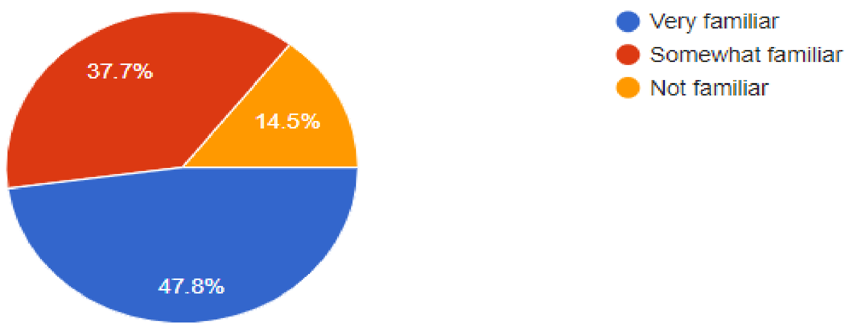


Figure 4.16. Analysis on Participant Familiarity with CPPs.

With an estimated rate of 47.8%, and 37.7%, majority of industry experts and manufacturing executives within 2-5 years and 5-10years category, responded to being "very familiar" which is a positive view and "somewhat familiar" respectively. However, a response rate of 14.5% was generated from participants that were completely unfamiliar with the concept of CPPs.

The next survey question was asked to determine the critical process areas of improvement using machine learning to streamline operations, the advantages it offers by detecting deviations or trigger corrective actions and its impact on decision making.

Q.16. How effectively do you believe machine learning could improve and control critical process parameters during manufacturing processes?

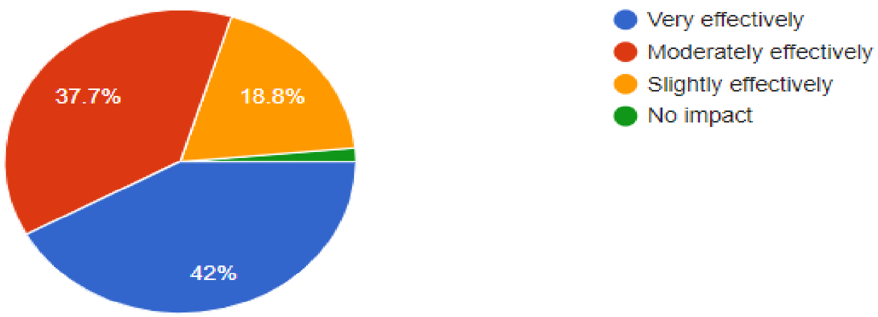


Figure 4.17. Analysis Rate on Participant's Opinion of ML to Monitor and Control CPP.

33 respondents were very confident in ML having a significant control over CPPs during manufacturing with a high rate of 42%. This category of participants is of the perception that ML can help optimize drug formulation by analysing the impact of different variables on product stability, bioavailability, and efficacy. A total of 26 respondents on the other hand believe that ML has a moderate effect yielding a rate of 37.7%. and 13 respondents were not entirely sure with a rate of 18.8%. The confidence level of majority of the respondents is on a higher which implies that has great potential to greatly improve and control CPPs in pharmaceutical manufacturing.

The next question in the survey aims to assess the operational factors relating to the manufacturing process that are target for overall process optimization. Answers were provided in checkboxes and participants could select more than one option.

Q.17. In your opinion what are the most important process parameters to consider in the manufacturing of drug dosage forms?

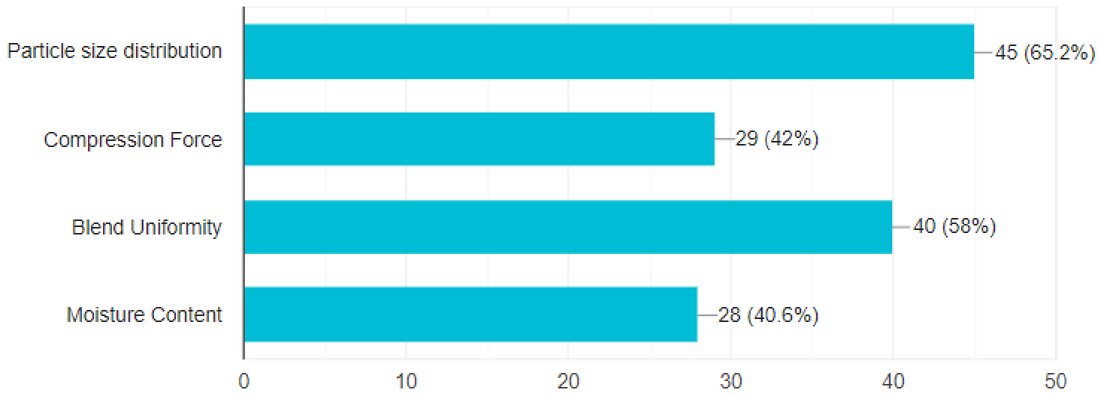


Figure 4.18. Analysis Rate of Process Parameters According to Participants.

The outcome of the analysis above indicates a high response rate of 65.2% with Particle size distribution having 45 response counts. This implies that uniform distribution of APIs and excipients can have an impact on achieving homogeneity and quality of final dosage form. 58% responses were obtained from 40 response counts on blend uniformity, 42% response rate on compression force with a total of 29 response counts and lastly 40.6% rate on moisture content with 28 response counts.

Q.18. What challenge(s) do you believe are the most prevalent during the monitoring and controlling of critical process parameters from raw material to finished product?

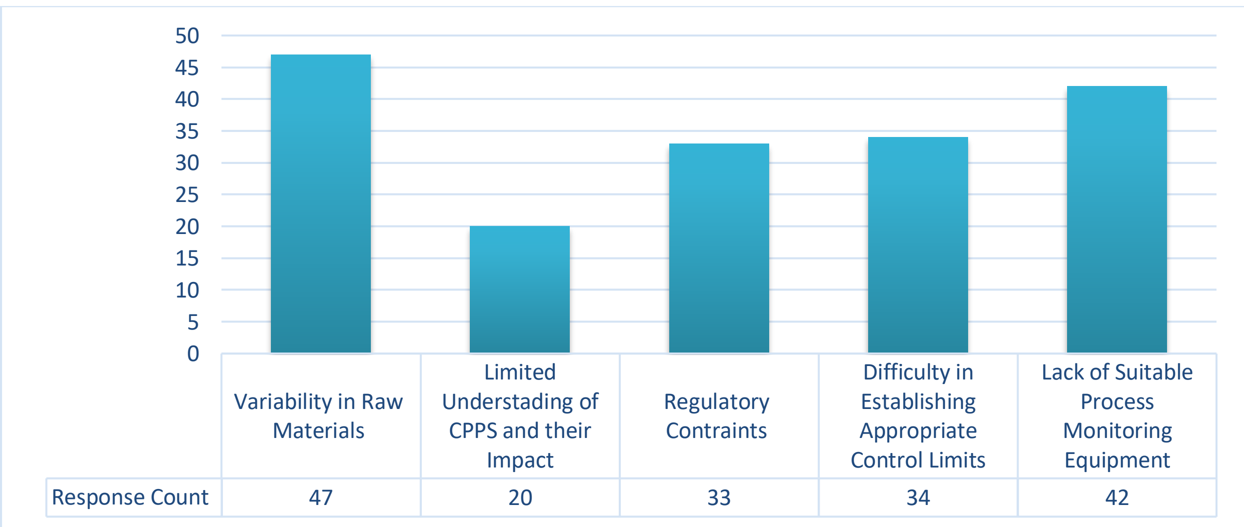


Figure 4.19. Analysis on Participant's Choice of Prevalent Challenges

As stated earlier, variability is a prevalent challenge commonly encountered in manufacturing dosage forms. However, there are other factors that could be brought about by the manufacturing equipment, workforce or regulatory requirement that can influence smooth manufacturing procedures. Figure 4.19 presents the analysis gotten from the respondents. Variability in raw materials had the highest response count of 47(68.1%), lack of suitable process monitoring equipment had a response count of 41 (60.9%), difficulty in establishing appropriate control limits had a response count of 34(49.3%), regulatory constraints had a response count of 33(47.8%) and limited understanding of CPPs and their impact had a response count of 20(29%). This implies that the challenges participants face is in direct correlation with the challenges mentioned in reviewed literature.

The next question aims to assess the confidence level of Participants on more precise control and optimization. We have learnt from literature the predictive capacity of ML algorithms to uncover patterns and relationships that are not immediately apparent to the workforce. By exploring participant's perception to this statement, the opinions from respondents would provide significant insights.

Q.19. How confident are you on the ability of machine learning to identify and optimise critical process parameters on the efficiency and cost-effectiveness of manufacturing processes?

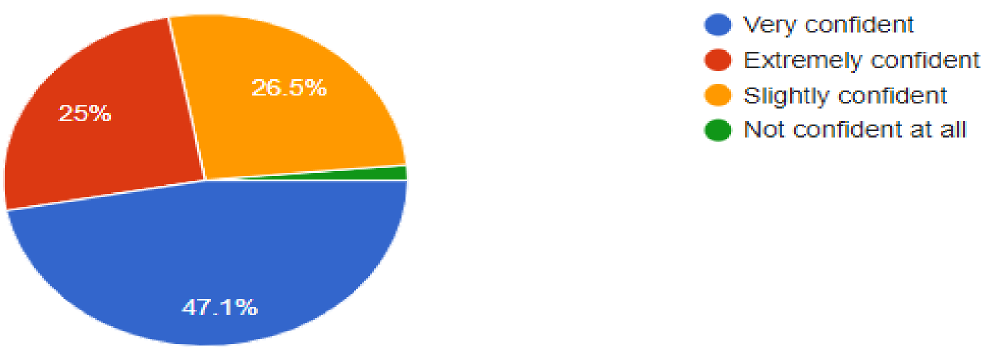


Figure 4.20 Analysis Rate on CPP optimization using ML According to Participants.

According to the chart, 32 respondents fully trust in ML abilities to adapt and make real-time adjustment to critical parameters ensuring consistent and high-quality production. This generated a response rate of 47.1%. 17 manufacturers and industry experts had a positive view and were extremely confident with a response rate of 25%., 18 respondent were slightly confident and a bit sceptical towards the statement with a response rate of 26.5% and 1 research specialist had no confidence in ML.

The last question of the survey aims to assess the tasks that are manually performed that could be automated with the implementation of machine learning and as a result, upskilling and reskilling may be required to adapt to algorithm development and collaboration with intelligent systems.

Q20. Please select your perceived opinion on the use of machine learning in the manufacturing procedures and its potential impact on job roles and responsibilities.

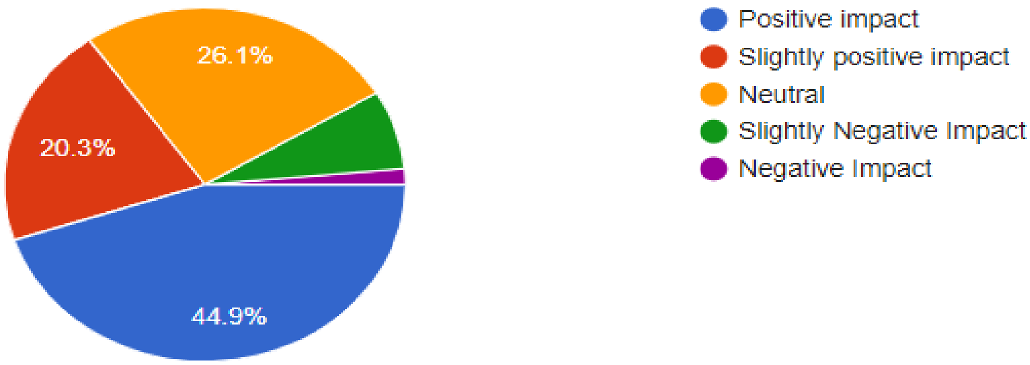


Figure 4.21. Analysis Rate on ML Impact According to Participants.

Majority of the respondents believed that use of machine learning would have a positive impact on manufacturing procedures with a response rate of 44.9% (32), 14 respondents were slightly confident with a response rate of 20.3%, 18 respondents were neural towards the statemen with a response rate of 26.1%.5 respondent slight believed that ML would have a negative impact with a response rate of 7.2% and there was a response rate of 1.4% from 1 respondent would is of a strong believe that ML has a negative impact on job roles in pharmaceutical manufacturing.

CHAPTER FIVE

5. Discussion and Conclusion

5.1 Research Summary

This is the main part of the research as it enables the readers to understand the results obtained in the study by evaluating key points, its implications in different areas and improvement that the study can make for future development (Taherdoost, 2022). This chapter discusses the theoretical importance in relation to the objectives of this thesis as mentioned. This chapter also encompasses a summary of the entire thesis, recommendations relating to the topic, and a reflective evaluation of the study. In today's rapidly evolving digitalization era, optimizing manufacturing processes is crucial for ensuring economic performance and maximising profits. Pharmaceutical manufacturing systems are inherently faced variability and high dimensionality, therefore, employing all available resources is crucial to meeting the demand for high quality pharmaceuticals and machine learning is a field that has seen rapid advancement in both usability and potential outcomes. Although, the broad expanse of the field poses challenge to its extensive adoption. This research aimed to address this by providing an overview of available ML techniques with emphasis on the potential benefits for process optimization and offers instances on successful applications in the manufacturing aspect.

As a reminder, the central research questions are;

- To what extent can the application of machine learning enhance operations of pharmaceutical manufacturing, what are the specific factors or emerging technologies that contribute to this AI-driven optimization approach and how do they differ across organizational context?
- What are the opportunities associated with the integration of machine learning in the pharmaceutical sector, in terms of quality attributes, uniformity, yield, and compliance with regulatory standards.
- How can machine learning be used in these manufacturing areas to assure public safety; what are the ethical and regulatory considerations.

The review of literature in PubMed, Google Scholar, and Research Gate on ML application for process optimization has formed the basis of the research questions and objective by providing a view of the current state of pharmaceutical manufacturing, existing applications of ML to various stages of process and product development, and tremendous opportunities to improve efficiency, accuracy, and predictability of pharmaceutical dosage formulation development process. It also revealed the expanding interest of ANNs and its substantial potentials in various pharmaceutical applications. A good example stated in literature by Nagaprasad *et al* (2021) was the recent use of ANNs to research and analyse the efficacy of CPPs which generated 13 critical predictions and evaluations. Primary research in this aspect also has revealed the participant's opinions on the types of ML algorithms used generally as the use of ML that have gained global population over the years specifically in the manufacturing sector. These participants were selected to gain a variety of opinions and views on the topic related to the research questions. A quantitative survey-questionnaire was distributed to 90 pharmaceutical and biopharmaceutical executives in drug formulation development to assess their value perception of machine learning in manufacturing processes and overall, 69 independent responses was received through the survey instrument. There were differences in the number of responses in each field and most responses came from process operators and production managers and quality control officers. It was evident from the first section of the survey, the opinions from 54 respondents who significantly indicated that they were aware of ML provided valuable insights for this study. According to this, most of the respondents either has a positive view regarding the concept and/or are currently using ML in the industry. The key theme was to understand whether the distribution of respondents on the different roles in pharmaceutical manufacturing industry highlights the widespread interest and potential application of ML for process optimization within different facilities.

To answer the second research question, the research conducted by Flach in 2022 considered the unique characteristics of different processes and variables during pharmaceutical manufacturing. It was revealed in his study, one of the advantages of ML algorithms provided, was the opportunity to gain experience from the dynamic system and to some extent, automate environmental adaptation. Depending on the choice of algorithm, the adaptation process can be faster than the traditional approaches which serves as the foundation for predicting the manufacturing system as utilising ML techniques can lead to the extraction of patterns from existing data sets. Based on this, the next section of the survey goes deep into context to gain better understanding of the participants' knowledge on the type of algorithms used to control process and product variability during manufacturing processes as well as opportunities associated with the integration and compliance with regulatory standards. Majority of the participants agreed with the phenomenon that the overall benefits of ML in pharmaceutical manufacturing is to achieve process efficiency and product quality where 43.5% were in line with regulatory compliance and predictive maintenance and 34.8% accounted for predictive maintenance. However, there seems to be a disconnection between understanding the advantages and value of ML and adopting this technology. One interesting example to back this up came from the second section of the survey which revealed that 63.2% of the participants with the impression of ML being moderately implemented were from respondents mostly in the field of industry experts and quality assurance specialist. This can be attributed to a few factors. One could be the perception that implementing ML requires significant investment in terms of resources, time, and expertise. Some organizations could also be hesitant due to regulatory concerns. Additionally, there might be lack of awareness about the specific benefits ML can bring to pharmaceutical manufacturing. It is important to bridge this gap by showcasing successful case studies and highlighting the positive impact ML can have on efficiency, quality and decision making in future research.

To answer the third research question, Mitchel (2013) revealed in his study the common CPPs that can impact the CQAs of the final product typically occurs in the initial phase of dosage formulation such as dispensing, granulating, tableting, coating, batch processing etc. To evaluate this, the third section of the survey further investigated the critical material attributes and quality parameters that significantly influence the critical quality attributes of the final product in real time and the areas of improvement using machine learning to streamline operations, the advantages it offers and its impact on decision making. One can say that the results of overall responses generated from the section showed that 54.6% were in agreement that ML in pharma has a positive impact on public health and this is in direct correlation with the reviewed literature. Emphasis is laid on the basic reality of implementing ML while taking into consideration the regulatory and ethical challenges that comes with it to assure public health as discussed in the next section.

5.2. Regulatory and Ethical Concerns Surrounding the Use of ML in Pharmaceutical Manufacturing.

ML is gradually finding its way in pharma and life science organizations as it has the potential to improve the safety of clinical trials and streamline the formulation development process in pharmaceutical manufacturing processes. However, it must be determined whether there are any risks or harm associated with its integration for instance, when applied outside of a controlled setting, ML has the capacity to magnify errors that already exists in data sources which raises ethical concerns. The following are some of the major ethical issues;

- Data bias needs to be avoided as pharmaceutical manufacturing involves several processes that rely would rely on using appropriate algorithm based on un-biased real time historical data.
- The research and application of machine learning in the pharmaceutical sector must be focused on achieving positive environmental outcomes rather than just maximising energy consumption, even though training the model requires substantial amount of energy. As such, to optimises machine learning's potential environmental benefits that aligns with process optimization, it should be developed and executed with a focus on sustainability and energy efficiency.
- Comprehending the inner workings of artificial intelligence systems becomes increasingly challenging as their complexity increases. A possible risk to public safety could result from this lack of transparency. Ensuring the clear and explicable architecture of machine learning systems is crucial.
- As machine learning technology can be applied to malicious endeavours like the development of autonomous weapons or hackings. The ethical application of ML and the implementation of suitable controls to prevent its misuse is imperative.
- Regular audits of the algorithms, including their integration into systems, and programming groups that are inclusive and varied are vital. All machine learning

predictions, including the quickest ones are methodical due to the use of ML algorithms, in contrast to human decision-making.

The use of ML in pharmaceutical manufacturing was considered with reference to public safety. Although ML can expedite the formulation development process, there are potential risks related to its application that needs to be carefully considered. In general, both literature and the survey instrument support the fact that professionals that were actively involved in pharmaceutical manufacturing see the advantages in the implementation of ML in drug formulation and development.

5.3. Conclusion

This thesis comprised five research chapters that examined the use of machine learning for process optimization in pharmaceutical manufacturing. This application considers machine learning as an applied technology to drive automation in specific areas of the pharmaceutical industry. Generally, large scale testing and multivariate data analysis are two statistical approaches that have been used traditionally for the identification and optimization of process variables. Artificial intelligence can offer an alternative by providing significant opportunities to increase the precision and effectiveness of drug dosage formulation, presented by the rapid development od machine learning algorithms. However, access to high quality data, resolving ethical concerns and understanding the limitations of ML-based methods are critical to its effective implementation.

It is important to adopt machine learning with the expectation that the end solution will enable automated, human like decision making, particularly for manufacturing improvements within the regulatory framework and design space for well-characterised manufacturing processes, as the study findings demonstrated effective continuous quality improvement in pharmaceutical manufacturing using this methodology. Overall, this research contributes to advancing the understanding of how machine learning algorithms can revolutionize process optimization and pave the way for the future of pharmaceutical manufacturing through insights, analysis, and evidence in the industry.

5.4 Limitation and Research Recommendations.

Due to time constraints, the percentage of individuals that participated in the research and those that denied taking part in the research resulted in a small sample size. The research was further limited by the data collection method, which involved using an online survey approach. This method may not have produced a sample size large enough for adequate statistical analysis. Therefore, more research is needed in the future to ensure that reliable data is transmitted between pharmaceutical manufacturing facilities. There is potential for more nuanced research into the specific applications of ML in various aspects of Research and Development, as this study provided a general overview of ML application in the manufacturing processes of drug dosage forms. This strategy could be useful in revealing the estimated cost and time to bring new drugs to the market. Furthermore, this could provide insightful information into a more precise measure of how ML is being used in R&D settings and how the pharmaceutical sector has changed over time in terms of adoption and use of the technology. It is also important to visit the areas ML application in pharma has not been successful. Most machine learning struggle to predict drug molecules' solubility with any degree of accuracy due to their complicated structure and numerous variables that influence it. Additionally, machine learning algorithms encounter difficulties when evaluating and predicting the behaviour of biological systems including protein folding or cell contact, in the production of biopharmaceuticals. There is a need for additional research to address these challenges in future**.**

REFERENCES

Ailisto, H., Helaakoski, H. and Neuvonen, A. (2023) (2023062224) DOI: 10.20944/preprints202306.2224.v1.

Alphonse, N. (2023) 'The Main Stages of the Research Process - A Review of the Literature'. *International Journal of Research and Review*, 10(7), pp. 671–675. DOI: 10.52403/ijrr.20230779.

Almanei, M. *et al.* (2021) 'Machine Learning Algorithms Comparison for Manufacturing Applications'. DOI: 10.3233/ATDE210065.

Alwis, D. *et al.* (2007) 'Statistical Methods in Media Optimization for Batch and Fed-Batch Animal Cell Culture'. *Bioprocess and Biosystems Engineering*, 30, pp. 107–13. DOI: 10.1007/s00449-006-0107-7.

Amin, M. and Ali, A. (2017) *Performance Evaluation of Supervised Machine Learning Classifiers for Predicting Healthcare Operational Decisions*. DOI: 10.13140/RG.2.2.26371.25127.

Arden, N.S. *et al.* (2021) 'Industry 4.0 for Pharmaceutical Manufacturing: Preparing for the Smart Factories of the Future'. *International Journal of Pharmaceutics*, 602, p. 120554. DOI: 10.1016/j.ijpharm.2021.120554.

Arifin, S.R. (2018) 'Ethical Considerations in Qualitative Study'. *INTERNATIONAL JOURNAL OF CARE SCHOLARS*, 1. DOI: 10.31436/ijcs.v1i2.82.

Asenahabi, B. (2019) 'Basics of Research Design: A Guide to Selecting Appropriate Research Design'. 6, pp. 76–89.

Awad, A. *et al.* (2021) 'Chapter 20 - Liquid Dosage Forms'. In Adejare, A. (ed.) *Remington (Twenty-Third Edition)*. Academic Press, pp. 359–379. DOI: 10.1016/B978-0-12-820007-0.00020-9.

Banner, M. *et al.* (2021) 'A Decade in Review: Use of Data Analytics within the Biopharmaceutical Sector'. *Current Opinion in Chemical Engineering*, 34, p. 100758. DOI: 10.1016/j.coche.2021.100758.

Bannigan, P. *et al.* (2021) 'Machine Learning Directed Drug Formulation Development'. *Advanced Drug Delivery Reviews*, 175, p. 113806. DOI: 10.1016/j.addr.2021.05.016.

Barmpalexis, P. *et al.* (2018) 'Development of a New Aprepitant Liquisolid Formulation with the Aid of Artificial Neural Networks and Genetic Programming'. *AAPS PharmSciTech*, 19(2), pp. 741–752. DOI: 10.1208/s12249-017-0893-z.

Bar-Or, A. *et al.* (2005) 'Decision Tree Induction in High Dimensional, Hierarchically Distributed Databases'. DOI: 10.1137/1.9781611972757.42.

Bautista, M. and Amigo, J. (2015) 'Manufacturing Semi-Solid and Liquid Dosage Forms: A Point of View from the NIR-PAT Perspective'. *Manufacturing Semi-Solid and Liquid Dosage Forms: A Point of View from the NIR-PAT Perspective*, 20, pp. 19–23.

Baviskar, K. *et al.* (2023) 'Artificial Intelligence and Machine Learning-Based Manufacturing and Drug Product Marketing'. In *Bioinformatics Tools for Pharmaceutical Drug Product Development*. John Wiley & Sons, Ltd, pp. 197–231. DOI: 10.1002/9781119865728.ch10.

Beg, S. *et al.* (2019) 'Chapter 3 - Application of Design of Experiments (DoE) in Pharmaceutical Product and Process Optimization'. In Beg, S. and Hasnain, M.S. (eds.) *Pharmaceutical Quality by Design*. Academic Press, pp. 43–64. DOI: 10.1016/B978-0-12-815799-2.00003-4.

Bhattacharyya, D.K. (2006) *Research Methodology*. Excel Books India.

Bhambure, R., Kumar, K. and Rathore, A.S. (2011) 'High-Throughput Process Development for Biopharmaceutical Drug Substances'. *Trends in Biotechnology*, 29(3), pp. 127–135. DOI: 10.1016/j.tibtech.2010.12.001

Bharat, P. and Paresh, M. (2023) 'A Review : Novel Advances in Semisolid Dosage Forms & Patented Technology in Semisolid Dosage Forms'.

Carleton, J.P.A., Frederick J. (ed.) (2013) *Validation of Pharmaceutical Processes*. 3rd Edition. Boca Raton: CRC Press DOI: 10.3109/9781420019797.

Chauhan, S. *et al.* (2021) (3) 'Using Peptidomics and Machine Learning to Assess Effects of Drying Processes on the Peptide Profile within a Functional Ingredient'. *Processes*, 9(3), p. 425. DOI: 10.3390/pr9030425.

Cheng, A. (2023) 'Process Optimization and Scale-up in Pharmaceutical Manufacturing.'

Chi, H.-M. *et al.* (2009) 'Machine Learning and Genetic Algorithms in Pharmaceutical Development and Manufacturing Processes'. *Decision Support Systems*, 48(1), pp. 69–80. DOI: 10.1016/j.dss.2009.06.010.

CHODANKAR, R. and Dev, A. (2016) 'Optimization Techniques: A Futuristic Approach for Formulating and Processing of Pharmaceuticals'. *Indian Journal of Pharmaceutical and Biological Research*, 4, pp. 32–40. DOI: 10.30750/ijpbr.4.2.5.

Chowdary, K.P.R. and Shankar, K.R. (2016) 'Optimization of Pharmaceutical Product Formulation by Factorial Designs:Case Studies'. *Journal of Pharmaceutical Research*, pp. 105–109. DOI: 10.18579/jpcrkc/2016/15/4/108815.

Coley, C.W., Eyke, N.S. and Jensen, K.F. (2020) 'Autonomous Discovery in the Chemical Sciences Part I: Progress'. *Angewandte Chemie International Edition*, 59(51), pp. 22858–22893. DOI: 10.1002/anie.201909987.

Colucci, D. *et al.* (2021) 'On-line product quality and process failure monitoring in freeze-drying of pharmaceutical products'. In *Drying Technology*. Drying Technology. Taylor & Francis, pp. 134–147. DOI: 10.1080/07373937.2019.1614949.

Cosenza, Z., Block, D.E. and Baar, K. (2021) 'Optimization of Muscle Cell Culture Media Using Nonlinear Design of Experiments'. *Biotechnology Journal*, 16(11), p. e2100228. DOI: 10.1002/biot.202100228.

Creswell, J.W. and Creswell, J.D. (2014) 'Research Design: Qualitative, Quantitative, and Mixed Methods Approaches'.

Cui, R. and Zhu, F. (2020) 'Ultrasound Modified Polysaccharides: A Review of Structure, Physicochemical Properties, Biological Activities and Food Applications'. *Trends in Food Science & Technology*, 107. DOI: 10.1016/j.tifs.2020.11.018.

Dahlgren, G. *et al.* (2019) 'Continuous Twin Screw Wet Granulation and Drying—Control Strategy for Drug Product Manufacturing'. *Journal of Pharmaceutical Sciences*, 108(11), pp. 3502–3514. DOI: 10.1016/j.xphs.2019.06.023.

Damiati, S. *et al.* (2017) 'Application of Machine Learning in Prediction of Hydrotrope-Enhanced Solubilisation of Indomethacin'. *International Journal of Pharmaceutics*, 530(1–2), pp. 99–106. DOI: 10.1016/j.ijpharm.2017.07.048.

Davidavičienė, V. (2018) 'Research Methodology: An Introduction'. In Marx Gómez, J. and Mouselli, S. (eds.) *Modernizing the Academic Teaching and Research Environment: Methodologies and Cases in Business Research*. Cham: Springer International Publishing, pp. 1–23. DOI: 10.1007/978-3-319-74173-4_1.

Diaz, L. *et al.* (2023) 'Machine Learning Approaches to the Prediction of Powder Flow Behaviour of Pharmaceutical Materials from Physical Properties'. *Digital Discovery*, 2. DOI: 10.1039/D2DD00106C.

Dibekulu, D. (2020) 'An Overview of Data Analysis and Interpretations in Research'. pp. 1–27. DOI: 10.14662/IJARER2020.015.

Djuris, J. *et al.* (2021) 'Application of Machine-Learning Algorithms for Better Understanding of Tableting Properties of Lactose Co-Processed with Lipid Excipients'. *Pharmaceutics*, 13(5), p. 663. DOI: 10.3390/pharmaceutics13050663.

Dong, J., Gao, H. and Ouyang, D. (2021) 'PharmSD: A Novel AI-Based Computational Platform for Solid Dispersion Formulation Design'. *International Journal of Pharmaceutics*, 604, p. 120705. DOI: 10.1016/j.ijpharm.2021.120705.

Dong, Y. *et al.* (2023) (7) 'Data-Driven Modeling Methods and Techniques for Pharmaceutical Processes'. *Processes*, 11(7), p. 2096. DOI: 10.3390/pr11072096.

Dörr, F. and Florence, A. (2020) 'A Micro-XRT Image Analysis and Machine Learning Methodology for the Characterisation of Multi-Particulate Capsule Formulations'. *International Journal of Pharmaceutics: X*, 2, p. 100041. DOI: 10.1016/j.ijpx.2020.100041.

Escotet-Espinoza, M.S., Rogers, A. and Ierapetritou, M.G. (2016) 'Optimization Methodologies for the Production of Pharmaceutical Products'. In Ierapetritou, M.G. and Ramachandran, R. (eds.) *Process Simulation and Data Modeling in Solid Oral Drug Development and Manufacture*. New York, NY: Springer, pp. 281–309. DOI: 10.1007/978-1-4939-2996-2_9.

Fahle, S., Prinz, C. and Kuhlenkötter, B. (2020) 'Systematic Review on Machine Learning (ML) Methods for Manufacturing Processes – Identifying Artificial Intelligence (AI) Methods for Field Application'. *Procedia CIRP*, 93, pp. 413–418. DOI: 10.1016/j.procir.2020.04.109.

Ficzere, M. *et al.* (2022) 'Real-Time Coating Thickness Measurement and Defect Recognition of Film Coated Tablets with Machine Vision and Deep Learning'. *International Journal of Pharmaceutics*, 623, p. 121957. DOI: 10.1016/j.ijpharm.2022.121957.

Fischer, R.-P. *et al.* (2024) 'Digital Patient Twins for Personalized Therapeutics and Pharmaceutical Manufacturing'. *Frontiers in Digital Health*, 5. DOI: 10.3389/fdgth.2023.1302338.

Flach, P. (2022) *Machine Learning: The Art and Science of Algorithms That Make Sense of Data*. Cambridge: Cambridge University Press DOI: 10.1017/CBO9780511973000.

Flores-García, E. *et al.* (2023) 'Beyond the Lab: Exploring the Socio-Technical Implications of Machine Learning in Biopharmaceutical Manufacturing'. In pp. 462–476. DOI: 10.1007/978-3-031-43670-3_32.

Floryanzia, S. *et al.* (2022) 'Disintegration Testing Augmented by Computer Vision Technology'. *International Journal of Pharmaceutics*, 619, p. 121668. DOI: 10.1016/j.ijpharm.2022.121668.

Galata, D.L. *et al.* (2021) 'Real-Time Release Testing of Dissolution Based on Surrogate Models Developed by Machine Learning Algorithms Using NIR Spectra, Compression Force and Particle Size Distribution as Input Data'. *International Journal of Pharmaceutics*, 597, p. 120338. DOI: 10.1016/j.ijpharm.2021.120338.

Geneva: World Health Organization. (2024) 'Benefits and Risks of Using Artificial Intelligence for Pharmaceutical Development and Delivery'.

Ghazwani, M. *et al.* (2023) 'Development of Advanced Model for Understanding the Behaviour of Drug Solubility in Green Solvents: Machine Learning Modelling for Small-Molecule API Solubility Prediction'. *Journal of Molecular Liquids*, 386, p. 122446. DOI: 10.1016/j.molliq.2023.122446.

Gibson, M. (ed.) (2013) *Pharmaceutical Preformulation and Formulation: A Practical Guide from Candidate Drug Selection to Commercial Dosage Form*. 2nd Edition. Boca Raton: CRC Press DOI: 10.3109/9781420073188.

Gómez Escudero, G. *et al.* (2021) 'MACHINE LEARNING IN THE FIELD OF MANUFACTURING'. *DYNA*, 96, pp. 600–604. DOI: 10.6036/10197.

Griffin, D.J. *et al.* (2023) 'Opportunities for Machine Learning and Artificial Intelligence to Advance Synthetic Drug Substance Process Development'. *Organic Process Research & Development*, 27(11), pp. 1868–1879. DOI: 10.1021/acs.oprd.3c00229.

Grzesik, P. and Warth, S.C. (2021) 'One-Time Optimization of Advanced T Cell Culture Media Using a Machine Learning Pipeline'. *Frontiers in Bioengineering and Biotechnology*, 9, p. 614324. DOI: 10.3389/fbioe.2021.614324.

Guerra, A.C. (2018) 'Machine Learning in Biopharmaceutical Manufacturing'. *European Pharmaceutical Review*, 23(4), pp. 62–65.

Han, R. *et al.* (2018) 'Predicting Oral Disintegrating Tablet Formulations by Neural Network Techniques'. *Asian Journal of Pharmaceutical Sciences*, 13(4), pp. 336–342. DOI: 10.1016/j.ajps.2018.01.003.

Han, R. *et al.* (2019) 'Predicting Physical Stability of Solid Dispersions by Machine Learning Techniques'. *Journal of Controlled Release*, 311–312, pp. 16–25. DOI: 10.1016/j.jconrel.2019.08.030.

Harding, J.A. and Kusiak, A. (2006) 'Data Mining in Manufacturing: A Review'. Available at: https://www.semanticscholar.org/paper/Data-Mining-in-Manufacturing%3A-a-Review-Harding-Kusiak/bb71a3354b80dda0293a67aa57df38cce97a2eaf (Accessed: 6 April 2024).

Hardy, I.J. and Cook, W.G. (2003) 'Predictive and Correlative Techniques for the Design, Optimisation and Manufacture of Solid Dosage Forms'. *Journal of Pharmacy and Pharmacology*, 55(1), pp. 3–18. DOI: 10.1111/j.2042-7158.2003.tb02428.x.

Hong, M.S. *et al.* (2018) 'Challenges and Opportunities in Biopharmaceutical Manufacturing Control'. *Computers & Chemical Engineering*, 110, pp. 106–114. DOI: 10.1016/j.compchemeng.2017.12.007.

Horr, A.M. (2022) 'Optimization of Manufacturing Processes Using ML-Assisted Hybrid Technique'. *Manufacturing Letters*, 31, pp. 24–27. DOI: 10.1016/j.mfglet.2021.10.001.

Ibrić, S. *et al.* (2012) (4) 'Artificial Neural Networks in Evaluation and Optimization of Modified Release Solid Dosage Forms'. *Pharmaceutics*, 4(4), pp. 531–550. DOI: 10.3390/pharmaceutics4040531.

Jariwala, N. *et al.* (2023) 'Intriguing of Pharmaceutical Product Development Processes with the Help of Artificial Intelligence and Deep/Machine Learning or Artificial Neural Network'. *Journal of Drug Delivery Science and Technology*, 87, p. 104751. DOI: 10.1016/j.jddst.2023.104751

Jiang, J. *et al.* (2022) 'Emerging Artificial Intelligence (AI) Technologies Used in the Development of Solid Dosage Forms'. *Pharmaceutics*, 14(11), p. 2257. DOI: 10.3390/pharmaceutics14112257.

Jie, Z., Tian, G. and Qu, H. (2023) 'Pharmaceutical Application of Process Understanding and Optimization Techniques: A Review on the Continuous Twin-Screw Wet Granulation'. *Biomedicines*, 11, p. 1923. DOI: 10.3390/biomedicines11071923.

Karde, H., Rathod, S. and Kohale, D. (2023) 'A Comprehensive Review on Pharmaceutical Dosage Form'. *International Journal of Advanced Research in Science, Communication and Technology*, pp. 285–291. DOI: 10.48175/IJARSCT-8696.

Kazemzadeh Farizhandi, A.A., Alishiri, M. and Lau, R. (2021) 'Machine Learning Approach for Carrier Surface Design in Carrier-Based Dry Powder Inhalation'. *Computers & Chemical Engineering*, 151, p. 107367. DOI: 10.1016/j.compchemeng.2021.107367.

Kazemi, P. *et al.* (2016) 'Computational Intelligence Modeling of Granule Size Distribution for Oscillating Milling'. *Powder Technology*, 301. DOI: 10.1016/j.powtec.2016.07.046.

Khan, M., Miller, A. and Heller, J. (2023) *Techniques and Manufacturing Applications in Machine- Learning*.

Kontoravdi, C., Samsatli, N.J. and Shah, N. (2013) ‘Development and Design of Bio-Pharmaceutical Processes’. *Current Opinion in Chemical Engineering*, 2(4), pp. 435–441. DOI: 10.1016/j.coche.2013.09.007.

Kumar, K., Panpalia, G. and Priyadarshini, S. (2011) ‘Application of Artificial Neural Networks in Optimizing the Fatty Alcohol Concentration in the Formulation of an O/W Emulsion’. *Acta Pharmaceutica*, 61(2), pp. 249–256. DOI: 10.2478/v10007-011-0013-7.

Kumar, S.A. *et al.* (2022) ‘Machine Learning and Deep Learning in Data-Driven Decision Making of Drug Discovery and Challenges in High-Quality Data Acquisition in the Pharmaceutical Industry’. *Future Medicinal Chemistry*, 14(4), pp. 245–270. DOI: 10.4155/fmc-2021-0243.

Khuat, T.T. *et al.* (2024) ‘Applications of Machine Learning in Antibody Discovery, Process Development, Manufacturing and Formulation: Current Trends, Challenges, and Opportunities’. *Computers & Chemical Engineering*, 182, p. 108585. DOI: 10.1016/j.compchemeng.2024.108585.

Lee, H. *et al.* (2022) ‘Deep Learning-Based Prediction of Physical Stability Considering Class Imbalance for Amorphous Solid Dispersions’. *Journal of Chemistry*, 2022, pp. 1–11. DOI: 10.1155/2022/4148443.

Lee, J. and Ha, S. (2009) ‘Recognizing Yield Patterns through Hybrid Applications of Machine Learning Techniques’. *Information Sciences*, 179, pp. 844–850. DOI: 10.1016/j.ins.2008.11.008.

Liu, W. *et al.* (2022) ‘Development and Validation of Machine Learning Models for Prediction of Nanomedicine Solubility in Supercritical Solvent for Advanced Pharmaceutical Manufacturing’. *Journal of Molecular Liquids*, 358, p. 119208. DOI: 10.1016/j.molliq.2022.119208.

Liu, X. *et al.* (2021) ‘Long Short-Term Memory Recurrent Neural Network for Pharmacokinetic-Pharmacodynamic Modeling’. *International Journal of Clinical Pharmacology and Therapeutics*, 59(2), pp. 138–146. DOI: 10.5414/CP203800.

Lou, H. *et al.* (2019) ‘The Application of Machine Learning Algorithms in Understanding the Effect of Core/Shell Technique on Improving Powder Compactability’. *International Journal of Pharmaceutics*, 555, pp. 368–379. DOI: 10.1016/j.ijpharm.2018.11.039.

Lou, H., Lian, B. and Hageman, M.J. (2021) ‘Applications of Machine Learning in Solid Oral Dosage Form Development’. *Journal of Pharmaceutical Sciences*, 110(9), pp. 3150–3165. DOI: 10.1016/j.xphs.2021.04.013.

Margulis, E. *et al.* (2021) 'Intense Bitterness of Molecules: Machine Learning for Expediting Drug Discovery'. *Computational and Structural Biotechnology Journal*, 19, pp. 568–576. DOI: 10.1016/j.csbj.2020.12.030.

Ma, X. *et al.* (2020) 'Application of Deep Learning Convolutional Neural Networks for Internal Tablet Defect Detection: High Accuracy, Throughput, and Adaptability'. *Journal of Pharmaceutical Sciences*, 109(4), pp. 1547–1557. DOI: 10.1016/j.xphs.2020.01.014.

Melnikovas, A. (2018) 'Towards an Explicit Research Methodology: Adapting Research Onion Model for Futures Studies'. *Journal of Futures Studies*, 23, pp. 29–44. DOI: 10.6531/JFS.201812_23(2).0003.

Mészáros, L. *et al.* (2022) 'UV/VIS Imaging-Based PAT Tool for Drug Particle Size Inspection in Intact Tablets Supported by Pattern Recognition Neural Networks'. *International Journal of Pharmaceutics*, 620, p. 121773. DOI: 10.1016/j.ijpharm.2022.121773.

Mishra, M. (2018) 'Semisolid Dosage Forms Manufacturing: Tools, Critical Process Parameters, Strategies, Optimization and Recent Advances'

Mitchell, M. (2013) 'Determining Criticality-Process Parameters and Quality Attributes Part I: Criticality as a Continuum'. *Biopharm International*, 26, pp. 38-+.

Mohammed, M. (2022) 'Machine Learning Approaches, Technologies, Recent Applications, Advantages and Challenges on Manufacturing and Industry 4.0 Applications'. *International Journal for Research in Applied Science and Engineering Technology*, 10, pp. 1114–1121. DOI: 10.22214/ijraset.2022.46362.

Mohan, B., Kamaraj, R. and Navyaja, K. (2022) 'Role of Artificial Intelligence in Parenteral Formulation: A Review'. *ECS Transactions*, 107(1), pp. 20013–20020. DOI: 10.1149/10701.20013ecst.

Mondal, P.P. *et al.* (2023) 'Review on Machine Learning-Based Bioprocess Optimization, Monitoring, and Control Systems'. *Bioresource Technology*, 370, p. 128523. DOI: 10.1016/j.biortech.2022.128523.

Monostori, L. (2003) 'AI and Machine Learning Techniques for Managing Complexity, Changes and Uncertainties in Manufacturing'. *Engineering Applications of Artificial Intelligence*, 16, pp. 277–291. DOI: 10.1016/S0952-1976(03)00078-2.

Muheem, A. *et al.* (2020) 'Recent Advances in the Development of Parenteral Dosage Forms'. In p. 382. DOI: 10.1201/9780367821678-5.

Munir, N. *et al.* (2021) 'Machine Learning for Process Monitoring and Control of Hot-Melt Extrusion: Current State of the Art and Future Directions'. *Pharmaceutics*, 13(9), p. 1432. DOI: 10.3390/pharmaceutics13091432.

Mukhamediev, R. *et al.* (2022) 'Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges'. *Mathematics*, 10, p. 2552. DOI: 10.3390/math10152552.

Muñiz Castro, B. *et al.* (2021) 'Machine Learning Predicts 3D Printing Performance of over 900 Drug Delivery Systems'. *Journal of Controlled Release*, 337, pp. 530–545. DOI: 10.1016/j.jconrel.2021.07.046.

Nagy, B. *et al.* (2022) 'Application of Artificial Neural Networks in the Process Analytical Technology of Pharmaceutical Manufacturing—a Review'. *The AAPS Journal*, 24(4), p. 74. DOI: 10.1208/s12248-022-00706-0.

Noorain. *et al.* (2023) 'Artificial Intelligence in Drug Formulation and Development: Applications and Future Prospects'. *Current Drug Metabolism*, 24(9), pp. 622–634. DOI: 10.2174/0113892002265786230921062205.

O'Mahony, M. *et al.* (2022) 'Enhancing the Capabilities of Fluid Bed Granulation through Process Automation and Digitalisation'.

Onuki, Y. *et al.* (2012) 'Contribution of the Physicochemical Properties of Active Pharmaceutical Ingredients to Tablet Properties Identified by Ensemble Artificial Neural Networks and Kohonen's Self-Organizing Maps'. *Journal of Pharmaceutical Sciences*, 101(7), pp. 2372–2381. DOI: 10.1002/jps.23134.

Ovuoraye, P. *et al.* (2023) 'Machine Learning Algorithm and Neural Network Architecture for Optimization of Pharmaceutical and Drug Manufacturing Industrial Effluent Treatment Using Activated Carbon Derived from Breadfruit (Treculia Africana)'. *Journal of Engineering and Applied Science*, p. 138. DOI: 10.1186/s44147-023-00307-4.

Pham, D. and Afify, A. (2005) 'Machine-Learning Techniques and Their Applications in Manufacturing'. *Proceedings of The Institution of Mechanical Engineers Part B-Journal of Engineering Manufacture - PROC INST MECH ENG B-J ENG MA*, 219, pp. 395–412. DOI: 10.1243/095440505X32274.

Polykovskiy, D. *et al.* (2018) 'Entangled Conditional Adversarial Autoencoder for de Novo Drug Discovery'. *Molecular Pharmaceutics*, 15. DOI: 10.1021/acs.molpharmaceut.8b00839.

Prasada, A. *et al.* (2020) ‘Application of Response Surface Methodology (RSM) in Statistical Optimization and Pharmaceutical Characterization of a Patient Compliance Effervescent Tablet Formulation of an Antiepileptic Drug Levetiracetam’. *Future Journal of Pharmaceutical Sciences*, 6, pp. 1–14. DOI: 10.1186/s43094-020-00096-0.

Quadrianto, N. and Buntine, W. (2017) ‘Regression’. In pp. 1075–1080. DOI: 10.1007/978-1-4899-7687-1_716.

Rai, R. *et al.* (2021) ‘Machine Learning in Manufacturing and Industry 4.0 Applications’. *International Journal of Production Research*, 59, pp. 4773–4778. DOI: 10.1080/00207543.2021.1956675.

Rantanen, J. and Khinast, J. (2015) ‘The Future of Pharmaceutical Manufacturing Sciences’. *Journal of Pharmaceutical Sciences*, 104(11), pp. 3612–3638. DOI: 10.1002/jps.24594.

Rathore, A.S. *et al.* (2023) ‘Artificial Intelligence and Machine Learning Applications in Biopharmaceutical Manufacturing’. *Trends in Biotechnology*, 41(4), pp. 497–510. DOI: 10.1016/j.tibtech.2022.08.007.

Ray, A. (2023) ‘Optimizing Bioprocessing Techniques for Pharmaceutical Manufacturing’.

Reddy, B.M. (2023) ‘Machine Learning for Drug Discovery and Manufacturing’. In Rai, B.K.Kumar, G.and Balyan, V. (eds.) *AI and Blockchain in Healthcare*. Singapore: Springer Nature, pp. 3–30. DOI: 10.1007/978-981-99-0377-1_1.

Reklaitis, G.V., Khinast, J. and Muzzio, F. (2010) ‘Pharmaceutical Engineering Science-New Approaches to Pharmaceutical Development and Manufacturing’. *Chemical Engineering Science*, 65, pp. iv–vii. DOI: 10.1016/j.ces.2010.08.041

Reutlinger, M. and Schneider, G. (2012) ‘Nonlinear Dimensionality Reduction and Mapping of Compound Libraries for Drug Discovery’. *Journal of Molecular Graphics & Modelling*, 34, pp. 108–117. DOI: 10.1016/j.jmgm.2011.12.006.

Rogers, A. and Ierapetritou, M. (2014) ‘Challenges and Opportunities in Pharmaceutical Manufacturing Modeling and Optimization’. In Eden, M.R.Siirola, J.D.and Towler, G.P. (eds.) *Computer Aided Chemical Engineering*. Proceedings of the 8 International Conference on Foundations of Computer-Aided Process Design. Elsevier, pp. 144–149. DOI: 10.1016/B978-0-444-63433-7.50015-8.

Säfsten, K. and Gustavsson, M. (2020) *Research Methodology : For Engineers and Other Problem-Solvers*. Studentlitteratur AB Available at: https://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-51270 (Accessed: 10 April 2024).

Sakthivel, M. *et al.* (2021) ‘Review about Semisolid Dosage Form’.

Salem, S. *et al.* (2022) (4019138) DOI: 10.2139/ssrn.4019138.

Sarkis, M. *et al.* (2021) (3) 'Emerging Challenges and Opportunities in Pharmaceutical Manufacturing and Distribution'. *Processes*, 9(3), p. 457. DOI: 10.3390/pr9030457.

Saunders, M. *et al.* (2019) '"Research Methods for Business Students" Chapter 4: Understanding Research Philosophy and Approaches to Theory Development'. In pp. 128–171.

Selekman, J.A. *et al.* (2017) 'High-Throughput Automation in Chemical Process Development'. *Annual Review of Chemical and Biomolecular Engineering*, 8(1), pp. 525–547. DOI: 10.1146/annurev-chembioeng-060816-101411.

Severson, K. *et al.* (2015) 'Elastic Net with Monte Carlo Sampling for Data-Based Modeling in Biopharmaceutical Manufacturing Facilities'. *Computers & Chemical Engineering*, 80, pp. 30–36. DOI: 10.1016/j.compchemeng.2015.05.006.

Sharp, M., ak, R. and Hedberg, T. (2018) 'A Survey of the Advancing Use and Development of Machine Learning in Smart Manufacturing'. *Journal of Manufacturing Systems*, 48. DOI: 10.1016/j.jmsy.2018.02.004.

Shelke, M.S., Waghmare, M.S. and Kamble, D.H. (2021) 'A REVIEW ON OPTIMIZATION TECHNIQUES IN PHARMACEUTICAL FORMULATION'. 9(11).

Shi, G. *et al.* (2021) 'Pharmaceutical Application of Multivariate Modelling Techniques: A Review on the Manufacturing of Tablets'. *RSC Advances*, 11, pp. 8323–8345. DOI: 10.1039/D0RA08030F.

Shruti, C.S.N. (2015) 'OPTIMIZATION TECHNIQUES: AN OVERVIEW FOR FORMULATION DEVELOPMENT'. Available at: https://www.academia.edu/37437358/OPTIMIZATION_TECHNIQUES_AN_OVERVIEW_FOR_FORMULATION_DEVELOPMENT (Accessed: 13 April 2024).

Shukla, J. (2017) 'Recent Advances in Semisolid Dosage Forms'. *IAIAR International Journal of Science, Engineering & Technology*, 1, pp. 1–33.

Simões, M.F. *et al.* (2020) 'Artificial Neural Networks Applied to Quality-by-Design: From Formulation Development to Clinical Outcome'. *European Journal of Pharmaceutics and Biopharmaceutics: Official Journal of Arbeitsgemeinschaft Fur Pharmazeutische Verfahrenstechnik e.V*, 152, pp. 282–295. DOI: 10.1016/j.ejpb.2020.05.012.

Singh, V. *et al.* (2009) 'Optimization of Actinomycin V Production by Streptomyces Triostinicus Using Artificial Neural Network and Genetic Algorithm'. *Applied Microbiology and Biotechnology*, 82(2), pp. 379–385. DOI: 10.1007/s00253-008-1828-0.

Singh, V. *et al.* (2016) 'Strategies for Fermentation Medium Optimization: An In-Depth Review'. *Frontiers in Microbiology*, 7, p. 2087. DOI: 10.3389/fmicb.2016.02087.

Sohail Arshad, M. *et al.* (2021) 'A Review of Emerging Technologies Enabling Improved Solid Oral Dosage Form Manufacturing and Processing'. *Advanced Drug Delivery Reviews*, 178, p. 113840. DOI: 10.1016/j.addr.2021.113840.

Steinwandter, V., Borchert, D. and Herwig, C. (2019) 'Data Science Tools and Applications on the Way to Pharma 4.0'. *Drug Discovery Today*, 24(9), pp. 1795–1805. DOI: 10.1016/j.drudis.2019.06.005.

Strachon, M. and Schongut, M. (2023) 'Application of Machine Learning for Predicting Bulk Behaviour of Active Pharmaceutical Ingredients'. *British Journal of Pharmacy*, 8. DOI: 10.5920/bjpharm.1385.

Stranzinger, S. *et al.* (2021) 'Review of Sensing Technologies for Measuring Powder Density Variations during Pharmaceutical Solid Dosage Form Manufacturing'. *TrAC Trends in Analytical Chemistry*, 135, p. 116147. DOI: 10.1016/j.trac.2020.116147.

Sun, Y. *et al.* (2013) 'Application of Artificial Neural Networks in the Design of Controlled Release Drug Delivery System'. *Advanced Drug Delivery Reviews*, 55, pp. 1201–15. DOI: 10.1016/S0169-409X(03)00119-4.

Szlęk, J. *et al.* (2022) (4) 'Puzzle out Machine Learning Model-Explaining Disintegration Process in ODTs'. *Pharmaceutics*, 14(4), p. 859. DOI: 10.3390/pharmaceutics14040859.

Taherdoost, H. (2022) 'How to Write an Effective Discussion in a Research Paper; a Guide to Writing the Discussion Section of a Research Article'. *Open Access Journal of Addiction and Psychology*, 5. DOI: 10.33552/OAJAP.2022.05.000609.

Taherdoost, H. (2016) 'Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research'. *International Journal of Academic Research in Management*, 5, pp. 18–27. DOI: 10.2139/ssrn.3205035.

Tengli, M. (2020) 'RESEARCH ONION: A SYSTEMATIC APPROACH TO DESIGNING RESEARCH METHODOLOGY'.

Tulsyan, A., Garvin, C. and Ündey, C. (2018) 'Advances in Industrial Biopharmaceutical Batch Process Monitoring: Machine-Learning Methods for Small Data Problems'. *Biotechnology and Bioengineering*, 115(8), pp. 1915–1924. DOI: 10.1002/bit.26605.

Tyagi, S. *et al.* (2023) (3) 'AI-Assisted Formulation Design for Improved Drug Delivery and Bioavailability'. *Pakistan Heart Journal*, 56(3), pp. 149–162.

Vamathevan, J. *et al.* (2019) 'Applications of Machine Learning in Drug Discovery and Development'. *Nature Reviews. Drug Discovery*, 18(6), pp. 463–477. DOI: 10.1038/s41573-019-0024-5.

Vishwakarma, D. *et al.* (2022) 'Optimization Technique in Pharmaceutical Formulation and Processing – Review Article'. 24.

Vora, L.K. *et al.* (2023) 'Artificial Intelligence in Pharmaceutical Technology and Drug Delivery Design'. *Pharmaceutics*, 15(7), p. 1916. DOI: 10.3390/pharmaceutics15071916.

Walsh, I. *et al.* (2022) 'Harnessing the Potential of Machine Learning for Advancing "Quality by Design" in Biomanufacturing'. *mAbs*, 14(1), p. 2013593. DOI: 10.1080/19420862.2021.2013593

Wang, W. *et al.* (2021) 'Computational Pharmaceutics - A New Paradigm of Drug Delivery'. *Journal of Controlled Release*, 338, pp. 119–136. DOI: 10.1016/j.jconrel.2021.08.030.

Westphal, E. and Seitz, H. (2021) 'A Machine Learning Method for Defect Detection and Visualization in Selective Laser Sintering Based on Convolutional Neural Networks'. *Additive Manufacturing*, 41, p. 101965. DOI: 10.1016/j.addma.2021.101965.

Widodo, A. and Yang, B.-S. (2007) 'Support Vector Machine in Machine Condition Monitoring and Fault Diagnosis'. *Mechanical Systems and Signal Processing*, 21, pp. 2560–2574. DOI: 10.1016/j.ymssp.2006.12.007.

Wu, Z. *et al.* (2021) 'Do We Need Different Machine Learning Algorithms for QSAR Modeling? A Comprehensive Assessment of 16 Machine Learning Algorithms on 14 QSAR Data Sets'. *Briefings in Bioinformatics*, 22(4), p. bbaa321. DOI: 10.1093/bib/bbaa321.

Wuest, T. *et al.* (2016) 'Machine Learning in Manufacturing: Advantages, Challenges, and Applications'. *Production & Manufacturing Research*, 4(1), pp. 23–45. DOI: 10.1080/21693277.2016.1192517.

Xi, H. *et al.* (2020) 'Characterization of Spray Dried Particles Through Microstructural Imaging'. *Journal of Pharmaceutical Sciences*, 109(11), pp. 3404–3412. DOI: 10.1016/j.xphs.2020.07.032.

Yang, X. *et al.* (2019) 'Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery'. *Chemical Reviews*, 119(18), pp. 10520–10594. DOI: 10.1021/acs.chemrev.8b00728.

Zangaro, F., Minner, S. and Battini, D. (2020) 'A Supervised Machine Learning Approach for the Optimisation of the Assembly Line Feeding Mode Selection'. *International Journal of Production Research*, 59, pp. 1–22. DOI: 10.1080/00207543.2020.1851793.

Zheng, P. *et al.* (2018) 'Smart Manufacturing Systems for Industry 4.0: Conceptual Framework, Scenarios, and Future Perspectives'. *Frontiers of Mechanical Engineering*, 13. DOI: 10.1007/s11465-018-0499-5.

Zhou, J. *et al.* (2020) 'Identifying Capsule Defect Based on an Improved Convolutional Neural Network'. *Shock and Vibration*, 2020, p. e8887723. DOI: 10.1155/2020/8887723.

GRIFFITH COLLEGE DUBLIN

APPENDIX

Survey Questionnaire

GRIFFITH COLLEGE

Dissertation title: Utilizing Artificial Intelligence (Machine Learning Algorithms) For Process Optimization in Pharmaceutical Manufacturing Processes

Dear Participant,

I am Arafat Johnson, a post-graduate student at Griffith College Dublin, studying Digital transformation (Life Science). I am conducting a research on the use of machine learning algorithms in pharmaceutical manufacturing and how it can potentially improve operations as the application of artificial intelligence (machine learning) encompasses a wide range of algorithms based on trained datasets that can be applied in several circumstances for overall process optimization.

I am happy to invite you to take part in this research study and grateful if you can complete this questionnaire. The confidentiality of the information provided will be maintained and utilised solely for the purpose of this dissertation.

Answering the survey will take about 10 minutes. Kindly answer the questions by selecting your preferred option. Thank you for your valuable time and participation.

Yours faithfully,

Arafatadelodun.johnson@student.griffith.ie

CONSENT *

By checking the agree box below, you agree to participate in this research voluntarily, having understood the nature of this project. The responses you provide will be anonymized and once submitted, they cannot be withdrawn as they will not contain any personal identifying information. By agreeing to take part to this research, you are not waving any of your legal rights.

[ ] I agree [ ] I disagree

SECTION 1- BACKGROUND & DEMOGRAPHICS

Traditionally, pharmaceutical manufacturing processes have relied on manual interventions, empirical adjustment, and predefined control strategies to optimize production efficiency and maintain product quality. While these methods have been effective to an extent, they lack the agility, adaptability and predictive capabilities required to address the dynamic nature of modern manufacturing environments.

As a professional actively involved in pharmaceutical manufacturing processes, I would like to know your opinions in this survey, on the use of artificial intelligence (machine learning) as a digital tool to optimize and improve process performance in your manufacturing facility.

GRIFFITH COLLEGE DUBLIN

Please fill in or tick ✓in the appropriate fields.

1. Which of the following best describes your current role of profession?
   [ ] Manufacturing/ Production Personnel (Process Operator, Process Engineer,/Technician, Production Manager)
   [ ] Quality Assurance/ Quality Control Specialist
   [ ] Industry Expert (Data Scientist R&D Professional, Operator Officer, Executive, Director)
   [ ] Research Specialist in academia (Scholar, Research Scientist)
2. Please choose the option that most accurately reflect your years of experience in the given role.
   [ ] Less than a year
   [ ] 2 – 5 years
   [ ] 5 – 10 years
   [ ] Above 10 years
3. How familiar are you with artificial intelligence (AI) and machine learning (ML) technologies?
   [ ] Very familiar
   [ ] Somewhat familiar
   [ ] Not familiar
4. Have you heard about the application of machine learning algorithms in pharmaceutical manufacturing processes prior to this survey?
   [ ] Yes
   [ ] No
5. To what extent do you believe that improving process optimization in pharmaceutical manufacturing may be accomplished using machine learning?
   [ ] Significantly improve
   [ ] Moderately improve
   [ ] Slightly improve
   [ ] No impact
6. Which of the following do you believe best describes the current adoption of machine learning technologies in pharmaceutical industries?
   [ ] Strongly implementing
   [ ] Moderately Implementing
   [ ] Not implementing

SECTION 2

In this section, questions on your knowledge, views, and perspectives on the application machine-learning based approach in pharmaceutical manufacturing would be asked.

7. What do you believe are the primary motivations for implementing machine learning technologies in pharmaceutical manufacturing (Please select all that apply)

   [ ] Enhance product quality

   [ ] Improve product efficiency

   [ ] Ensure regulatory compliance

   [ ] Enable predictive maintenance

8. Have you implemented any machine learning-based approaches to optimise processes in pharmaceutical manufacturing?

   [ ] Yes

   [ ] No

9. If yes, kindly specify the type(s) of algorithm used.

   [ ] Predictive modelling (artificial neural network, support vector machines, decision tree, random forest

   [ ] Anomaly detection (Dimension reduction/ Principal component analysis, Clustering analysis)

   [ ] Other ____

10. Which aspect of pharmaceutical manufacturing processes do you believe can benefit the most from machine learning-based optimization? (Please select all that apply).

   [ ] Raw material selection

   [ ] Batch processing

   [ ] Quality Control

   [ ] Equipment maintenance

   [ ] Supply Chain Management

11. How would you rate the availability and quality of data required for machine learning-based process optimization based on your experience?

   [ ] Excellent

   [ ] Good

   [ ] Fair

   [ ] Poor

12. Do you believe that the use of machine learning in pharmaceutical manufacturing can assure public safety of products?

   [ ] Strongly agree

   [ ] Agree

   [ ] Neutral

   [ ] Disagree

   [ ] Strongly Disagree

13. Which enabling technologies do you believe contributes to the use of machine learning? (Please select all that apply).

[ ] Big data

[ ] Robotics and automation

[ ] Internet of things

[ ] Data Collection and Monitoring Systems (Sensors and Probes)

[ ] Process Analytical Technology

14. Which do you believe to be the likely outcome of the use of machine learning in manufacturing processes?

	Likely	Unlikely
Reduced Cycle time, variability & resource consumption.	[ ]	[ ]
Optimized formulation Stability, bioavailability, and efficacy	[ ]	[ ]
Enhanced Drug Release Profile Prediction	[ ]	[ ]

SECTION 3

In this section, questions on crucial processing factors that are essential to the manufacturing plant would be asked. This would help to examine the current challenges influencing the process and product lifecycle and how machine learning can improve these processes.

15. How familiar are you with the concept of critical process parameters in pharmaceutical manufacturing?

[ ] Very familiar

[ ] Somewhat familiar

[ ] Not familiar

16. How effectively do you believe machine learning could improve and control critical process parameters during manufacturing process?

[ ] Very effectively

[ ] Moderately effectively

[ ] Slightly effectively

[ ] No impact

17. In your opinion, what are the most important process parameters to consider in the manufacturing of drug dosage forms?

[ ] Particle size distribution

[ ] Compression Force

[ ] Blend Uniformity

[ ] Moisture Content

18. What challenge(s) do you believe are the most prevalent during the monitoring and controlling of critical process parameters from raw materials to finished product? (Please select all that apply).
    [ ] Lack of suitable process monitoring equipment
    [ ] Difficulty in establishing appropriate control limits
    [ ] Variability in raw materials
    [ ] Limited understanding of CPPs and their impact
    [ ] Regulatory Constraints
19. How confident are you in the ability of machine learning algorithms to identify and optimise critical process parameters on the efficiency and cost-effectiveness of manufacturing processes?
    [ ] Very confident
    [ ] Extremely confident
    [ ] Slightly confident
    [ ] Not confident at all
20. Please select your perceived opinion on the use of machine learning in the manufacturing procedures and its potential impact on job roles and responsibilities.
    [ ] Positive impact
    [ ] Slightly positive impact
    [ ] Neutral
    [ ] Slightly Negative Impact
    [ ] Negative Impact