International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 02 | February - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

MACHINE LEARNING CONTRIBUTIONS IN BIOPHARMACEUTICALS

Bhavna Venkatakrishnan, Swadha Singh
R.V College of Engineering, B.E Final Year Students

ABSTRACT

Machine learning (ML) has made significant contributions to the biopharmaceutical field, and its applications are increasing in quality, design-based development and manufacturing of biologics, reflecting the huge potential of bioprocess automation.We are still in the early stages in terms of direct support.This hinders the process from development to production. However, the accumulation of extensive production data has led to a significant increase in the use of ML-based models instead of traditional multivariate data analysis techniques. This trend is primarily driven by the introduction of advanced process analytical techniques to monitor biopharmaceutical process variables and quality attributes in real-time.To create accurate, flexible, and powerful predictive models that can address analysis, monitoring, and control problems in the biopharmaceutical field, taking into account the complexity and multidimensionality of bioproduct design, bioprocess development, and product manufacturing data.

ML-based approaches are increasingly being used. The purpose of this article is to provide a comprehensive overview of current applications of ML solutions in the design, monitoring, control and optimization of monoclonal antibody upstream, downstream and product formulation processes.Finally, this article details the key challenges associated with the use of the bioprocess itself, process data, and machine learning models in the development and manufacturing of monoclonal antibody processes.

Keywords
Biopharmaceuticals,Machine learning,UpstreamDownstream,Bioprocesses.

1.INTRODUCTION

In recent years, biopharmaceuticals, also called biologics, such as monoclonal antibodies (mAbs) and therapeutic proteins, have become the best-selling drugs in the pharmaceutical market (Lu et al., 2020), with seven of them expected to be sold in 2021.It has been the best-selling medicines around the world were biologics (Urquhart, 2022.

As defined by the U.S. Food and Drug Administration (FDA) (2018), biological products include vaccines, monoclonal antibodies, blood and blood components, allergens, somatic cells, tissues, gene therapies, and other therapeutic products containing protein.

Unlike small molecule or chemically synthesized drugs, biologics are complex and large mixtures of sugars, proteins, nucleic acids, and other substances that are difficult to easily identify or accurately characterize (Peters and Hennessey, 2020). Most biologics are produced in living systems such as microbial, plant, animal, or human cells, including mammalian cells such as Chinese hamster ovary cells (CHO), mouse myeloma (NS0), and baby hamster kidney (BHK) Produced by biotechnology.

Typically human embryonic kidney cells (HEK-293) and human retinal cells are used (Wurm, 2004). The biopharmaceutical market has experienced explosive growth in recent years.Since 2014, new biologics have been approved by the FDA every year for the treatment of various human diseases, including cancer, autoimmune diseases, metabolic diseases, and infectious diseases (De la Torre and Albericio, 2022) (see Figure 2 for details). Among biological products, mAbs have proven to be key products in the rapidly growing high-quality biologics market (Papathanasiou et al., 2019).

According to a report by Grand View Research (2022), the global therapeutic monoclonal antibodies market was valued at approximately $185.5 billion in 2021 and is expected to achieve a compound annual growth

rate of 11.30% from 2022 to 2030.This equates to approximately $494 billion in sales. If achieving this goal requires higher process productivity, more affordable end products, shorter production times, etc., lower cost solutions are available without sacrificing process robustness and product quality and intelligent manufacturing (Xenopoulos, 2015).

The fabricating handle of mAbs is complex with numerous interconnected components. It starts with the culture of hereditarily adjusted cells in little- and medium-sized bioreactors and growing them to the huge generation bioreactors inside upstream forms. After that, different downstream refinement steps are performed to isolate the item of interest from cell-derived and process-related debasements (Papathanasiou and Kontoravdi, 2020). Some time recently the Quality by Plan (QbD) has been presented by FDA as a standard procedure (Rathore and Winkle, 2009), the confirmation of last item quality within the ordinary fabricating of mAbs has more often than not depended on exorbitant and time-consuming quality by test strategies (Zhang and Mao, 2017). Within the quality by test approaches, the root causes of disappointment and/or critical levels of instability inside bioprocesses are frequently not well caught on. Consequently, it is as often as possible time-consuming, exorbitant, and possibly adversely affecting item security when the fabricating strategies ought to be rehashed until the root causes of disappointment are recognized and settled. In differentiate to the quality by test, QbD strategies are based on the all encompassing understanding of products and how materials and basic prepare parameters (CPPs) impact the basic quality traits (CQAs) and surrender of the ultimate items by means of distinctive unit operations (Zhang and Mao, 2017).In the QbD system, the item determinations, moreover known as the quality target item profiles (QTPPs), are set at the heart of the plan techniques of tests to ensure the quality and consistency of the ultimate items (Papathanasiou and Kontoravdi, 2020). This leads to the significant parts of in-process explanatory estimations and handling expository innovation (PAT). Besides, it is carved to get the relationship between handle conditions and item quality. These characteristics give openings for the utilization of computational models as promising options to test approaches to optimize and control the method practices more efficiently (Smiatek et al., 2020), to extricate valuable information from procured test information and to construct prescient models for in-silico tests. As displayed in Papathanasiou and Kontoravdi (2020), there are three fundamental challenging zones within the implementation of QbD standards in biomanufacturing:

(i) item heterogeneity, (ii) different confinements of estimations for handling checking, and (iii) the need of online preparation control procedures. These challenges can be moderated by taking advantage of knowledge-driven and data-driven approaches for (real-time) bioprocess checking and control based on the collected bioprocess information.

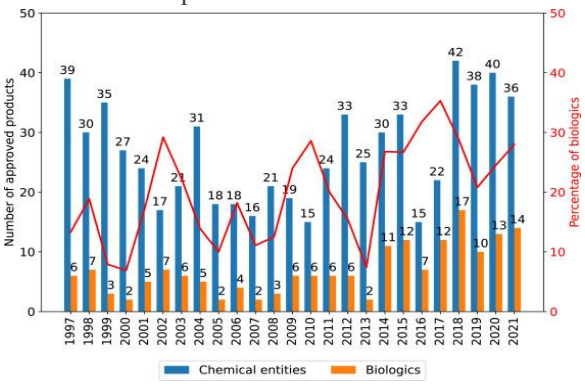

Fig. 2. New drugs have been approved by FDA over the last 25 years and the percentage of biologics over total drug approved each year.

By observing the connections among factors and extracting important information from bioprocess information, machine learning (ML) models can secure novel experiences into the interdependency between CPPs and CQAs in biopharmaceutical preparation advancement and fabricating. These experiences have the potential to help within the improvement of more viable preparation control techniques. Separated from real-time observing and control of upstream and downstream bioprocesses, administered ML calculations can be connected to disclosure and plan of biopharmaceuticals such as mAbs determination, media screening, plan of bolstering procedures, and choice of substrate for purification. ML calculations can too be utilized in product formulation and soundness steps. As a result, the sending of ML strategies for bioprocess advancement and fabricating can contribute to decreasing the time and generation fetched of mAbs and, in turn, bringing down the cost of counter acting agent treatment (Puranik et al., 2022, Rathore and Singh, 2015). Over the past few years, there has been watched a noteworthy increment within the utilize of multivariate information examination (MVDA), which is frequently considered to be a sub-field of ML centering on a bunch of measurable approaches, as a

PAT instrument for cell development (Pennant et al., 2021, Rathore and Singh, 2015) to investigate the accessible information from pilot or fabricating plants pointing to discover the relationship between handle factors and item quality properties, upgrade the bioprocess comprehension, and online observing and control of unit operations. As abridged in Mercier et al. (2014), well known applications of MVDA consist of the analysis of spectroscopic estimations and information profiles of unit operations like cell culture and chromatography, quantitative assessment of handle comparability, root cause investigation of disappointments, and crude fabric choice. MVDA strategies have been suggested by controllers as a basic device for QbD and PAT systems to make strides understanding of bioprocesses and to extend the levels of handle checking and control and, in turn,increasing the probability of accomplishing the required last item quality (Mercier et al., 2014).

As appeared in Standard et al. (2021), it ought to not be shocking on the off chance that the use of ML algorithms within biopharmaceuticals would develop dangerously in coming years to require benefits from investigated experiences and the control of predictive models to enhance, screen, and control bioprocess improvement and fabricating operations. The most patterns can center on the utilize of progressed information analytics, the integration of delicate sensors and PAT approaches to define progressed control techniques (Mandenius and Gustavsson, 2014), and the improvement of cross breed data-driven and knowledge-driven models related with coordinates plans of generation components such as handle, item, and cells (Badr and Sugiyama, 2020, Walsh et al., 2022). Besides, the biopharma field might witness a quick move to advanced transformation with the increased selection of queryable and organized centralized repositories such as a information lake or information distribution center for recorded information (Steinwandter et al., 2019). This digitisation would permit the application of ML models to all information assembled from numerous destinations over different scales and unit operations of manufacturing plants much appreciated to information consolidation and availability (Banner et al., 2021). This makes an imperative premise for the completely computerized data-driven biopharmaceutical fabricating offices prepared with information of physicochemical properties of substances, bioprocesses, and items in the near future.

The improvement and fabricating forms of biologics include an assorted run of bioproducts, including mAbs, antibodies, cell and gene treatments, and other helpful proteins, each with distinct production processes. This audit particularly centers on counter acting agent forms to demonstrate the common application of machine learning in biopharmaceutical forms. As a result, the paper points to a comprehensive survey of the existing applications of ML to different stages of process advancement and fabricating for mAbs, recognize challenges with respect to bioprocesses themselves and handle information, at that point proposing potential inquiries about bearings in this range. We don't point to supply measurements of the numbers of calculations used for specific problem spaces year over year. The interested readers can find more data for this topic in later surveys (Pennant et al., 2021, Guerra and Glassey, 2018, Mowbray et al., 2021, Pham et al., 2022). Instep, existing trends regarding how ML calculations are utilized within the handle advancement and manufacturing of biological products over diverse unit operations are checked on in detail. This paper is an expansion of the later audit paper (Puranik et al., 2022), which covers as it were commonplace applications. On the contrary, this paper presents a much bigger number of applications of ML to biopharmaceuticals and identifies numerous potential headings for the applications of progressed ML strategies to address existing challenges within bioprocess information pointing to the development of digital twin frameworks for biopharma 4.0. In this survey, we are going to explore the application of ML models in tending to challenges inside the advancement and fabricating of mAbs. Be that as it may, we will not dive into the coordinate comparison of calculation execution or offer particular calculation proposals. Assessing calculation execution requires a comprehensive benchmarking ponder with thorough comparison beneath differing criteria custom fitted to particular issues. For occasion, a benchmarking ponder (Poth et al., 2022) examining 18 ML models for real-time Raman-based expectations of key analytes in mammalian cell culture bioreactors revealed neural systems and irregular woodland relapse as optimal arrangements for this specific challenge. It is worth noticing that giving a broad depiction of different ML strategies surpasses the scope of this paper. Instep, we as it were to count the ML models and utilize tables to abridge which strategy is connected to each particular problem. For users interested in more point by point calculations, we

suggest alluding to textbooks such as Goodfellow et al. (2016) for profound learning and Hastie et al. (2009) for conventional ML calculations.

The rest of this paper is sorted out as takes after:

Area 2 abridges a few foundations of a typical biopharmaceutical handle for manufacturing mAbs. Following, Area 3 presents challenges regarding bioprocesses and process information characteristics, which can cause difficulties in building ML models. After that, Area 4 audits well known applications of ML in early stages of counter acting agent development, while Area 5 covers diverse applications of ML models in upstream preparing forms. Segment 6 is devoted to showing ordinary applications of ML within the downstream decontamination handle. The utilization of ML calculations to bargain with problems in item formulation and steadiness is displayed in Area 7. Concluding comments of this paper are depicted in Segment 9.

2) A typical bioprocess for production of monoclonal antibodies

The standard fabricating and refinement handle of mAbs incorporates two fundamental stages:

upstream (USP) and downstream handling (DSP) in combination with the ultimate item detailing and solidness, fill and wrap up stages. The upstream prepare comprises the development of cells in a arrangement of bioreactors and the generation of the interested item, whereas the downstream handle incorporates a arrangement of separation/purification steps to capture the item of intrigued and dispose of different handle, have cell, and item related debasements with negligible abdicate misfortune (Shukla and Thommes, 2010, Hong et al., 2018, Papathanasiou et al., 2019, Badr and Sugiyama, 2020, Papathanasiou and Kontoravdi, 2020). As of now, these operations are performed in fed-batch development frameworks and group partition forms (Xenopoulos, 2015), and now and then nonstop development frameworks to utilize these strategies. Fig. 3 speaks to a stream chart for a normal fabricating handle stage of mAbs and recombinant proteins.

The upstream handle starts with a vial defrost taken after by the inoculum development of cells from a cell bank and appropriate media, which is performed in shake jars or spinner carafes. After that, the upstream handle formally begins with the cell culture in seed bioreactors with the dynamic increment in measure and/or volume. The cells are at that point exchanged to the generation bioreactor for protein expression within the medium (Shukla and Thommes, 2010, Papathanasiou et al., 2019). At long last, a normal gather method is conducted to kill cells and cell flotsam and jetsam by centrifugation taken after by profundity and film filtration earlier to be transferred to decontamination steps within the downstream (Shukla and Kandula, 2008). Fed-batch approach is ordinarily utilized for the upstream preparation, where little volumes of nourish supplements are included to the bioreactors amid the cell development stages. Separated from the checking of nutrient and metabolite concentration, other variables of bioreactors within the upstream preparing are too regularly controlled amid the cell development counting broken down oxygen (pO2), pH, temperature and mass exchange of oxygen and CO2 (Shukla and Thommes, 2010).

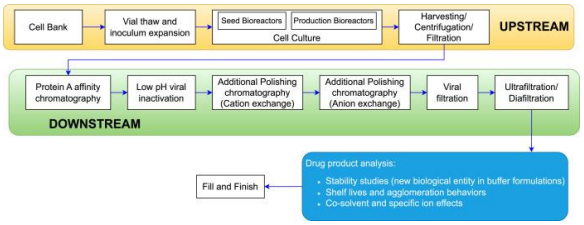

Fig. 3. Production process of mAbs.

Adapted from Papathanasiou et al., 2019, Smiatek et al., 2020, Shukla and Thommes, 2010, Hong et al., 2018, Badr and Sugiyama, 2020.

The downstream preparation for the fabricating of mAbs begins with Protein A chromatographic capture. This sort of fondness chromatography depends on the particular authoritative intuitive between the Fc zone of mAbs and the Protein A ligand. As demonstrated in Shukla and Thommes, 2010, Papathanasiou et al., 2019, Protein A liking chromatography step can lead to over 98% virtue since of tall official fondness and specificity of the Protein A to Fc-fusion proteins. This specificity empowers pollutants such as cell proteins, DNA and other contaminants to pass through whereas the items of intrigued ties to the stationary stage. The items are at that point eluted from Protein A liking adsorbents beneath moo pH conditions for infection

inactivation and decrease of the fondness of Protein A to the Fc districts. Another, one or numerous cleaning chromatographic steps are conducted to encourage evacuate handle and item related pollutions, which cannot be expelled by the Protein A fondness chromatography such as totals and item variations since of their chemical similitude with the determined protein as well as a byproduct related to the method like filtered protein A. Cation trade (CEX) chromatography and anion exchange (AEX) chromatography are the foremost commonly utilized cleaning approaches (Shukla and Thommes, 2010), which may be utilized independently or in combination. After cleaning steps, a viral filtration step is conveyed for infection expulsion. The ultimate step within the downstream prepare is regularly an ultrafiltration/diafiltration (UF/DF) to diminish capacity volumes and pass the item through the buffer for the definition of the sedate substance.The another organize of the fabricating handle of mAbs centers on advancing and dissecting the helpful items. Fundamental steps of this arrangement incorporate the considerations with respect to the plan of item details guaranteeing long shelf-lives and tall steadiness of the mAbs within the buffer arrangement in line with the required conveyance strategy (Smiatek et al., 2020). The fill and wrap up preparation constitutes the ultimate organization within the fabricating of mAbs and other biopharmaceuticals. This stage includes the exchange of the sterile sedate substance, produced and filtered in prior fabricating steps, into its last holder (vials, syringes, etc.), taken after by sealing for conveyance and utilization (Algorri et al., 2022). Within the majority of occasions, the fill and wrap up preparation for mAbs fabrication is robotized, utilizing adaptable aseptic automated fill wrap up offices (Dreckmann et al., 2020, Algorri et al., 2022). Computerization guarantees the last mAb product's reasonability, consistency over item clumps, and traceability from a Great Fabricating Hone (GMP) perspective. Moreover, it empowers the recording and following of all prepared parameters in compliance with regulatory prerequisites, guarantees exact volume, and decreases the probability of human mistakes (Sethi and Cunningham, 2021). The fill and wrap up operations are conducted utilizing advanced and exact filling machines, which are computerized, so there's as of now no writing accessible that applies ML calculations particularly to this stage. As a result, this survey paper will not envelop the fill and wrap up steps.Inside the QbD scheme, each biopharmaceutical manufacturing process has got to comply with predefined guidelines to guarantee the QTPPs of the ultimate item. In specific, the QbD strategy presents a set of CQAs which display the appropriate constraint, run, or conveyance of physical, chemical, natural, or microbiological properties or characteristics of in-process and wrapped up items (Mitchell, 2013). As claimed in Eon-Duval et al. (2012), these CQAs are habitually chosen based on risk-based examinations, where potential or known item quality properties are surveyed for their potential impacts on quiet security and product efficacy. Quality properties of recombinant restorative proteins can be separated into three bunches, i.e., product-related pollutions and substances, process-related pollutions, and contaminants (Eon-Duval et al., 2012, Papathanasiou et al., 2019). These commonly utilized properties are displayed in Table 1.This section only abridges the main steps and their usefulness inside a normal fabricating stage of mAbs. The survey of affecting components and the way of improving efficiency and altering item quality in conjunction with advancement techniques can be found in Radhakrishnan et al. (2018) for the upstream and in Baumann and Hubbuch, 2017, Rathore and Kapoor, 2014 for the downstream forms. Concurring to Badr and Sugiyama (2020), the advancement activities of the upstream development forms are as a rule conducted aiming to raise item amount represented by titer and productivity, whereas the optimisation of the downstream decontamination forms is performed to get the required item profile.

Table 1. Critical quality attributes are commonly used to assess the final recombinant therapeutic proteins.

Product-related impurities and substances	**Process-related impurities**	**Contaminants**
Aggregation	Residual host cell proteins (HCPs)	Adventitious agents (e.g. potential virus, bacteria, mycoplasma, bioburden, and fungi)
Fragmentation	Residual DNA	Endotoxins

N/C-terminal modifications	Raw material-created impurities (e.g. leached protein A, cell culture media, and buffers)
Oxidation	Leachables (from product contact materials)
Deamidation/Isomerization	
N-linked Glycosylation	
O-linked Glycosylation	
Glycation	
Conformation	
Disulfide bond modifications/free thiols	
GlcNAc, N-acetylglucosamine	

Adapted from Eon-Duval et al., 2012, Papathanasiou et al., 2019, Chan, 2019.

The fabricating handle of the mAbs is related with tight directions characterizing the immaculateness and composition of the ultimate items and the impediment of item and process-related pollution (Papathanasiou et al., 2019). Subsequently, thorough preparation checking gets to be basic for the fabricating handle. In that respect, control methodologies for following the real-time values of pollution and auxiliary items may be of extraordinary potential and tall request. These are the openings for the sending of data-driven modeling and enhancement for the entire biopharmaceutical fabricating preparation from the revelation and plan stages to upstream, downstream, and item detailing and steadiness. This paper will audit key ranges where ML algorithms can be utilized for these purposes. Be that as it may, to begin with we are going to examine challenges related with forms and bioprocess information.

3)Challenges of biopharmaceutical processes with respect to ML solutions

The biopharmaceutical industry stands at the precipice of a paradigm shift. Machine learning (ML) solutions offer immense potential to streamline processes, optimize production, and accelerate drug discovery. However, integrating ML into complex biological systems presents unique challenges that demand careful consideration and innovative solutions. This essay delves into the intricate landscape of biopharmaceutical processes, dissecting the critical hurdles that hinder the seamless adoption of ML solutions.

Data Scarcity and Heterogeneity:Biopharmaceutical processes generate intricate data, often riddled with sparsity and inconsistency. Samples are limited, experiments are expensive, and batch-to-batch variability introduces noise into the data. This poses a significant challenge for ML models, which require large, consistent datasets to learn effectively. Techniques like transfer learning and data augmentation can mitigate this issue, but their effectiveness relies heavily on domain knowledge and careful implementation.

Interpretability and Explainability:The "black box" nature of many ML models presents a crucial roadblock in biopharmaceutical applications. Regulatory agencies demand clear understanding of

how models arrive at their predictions, especially for decisions impacting patient safety and drug efficacy. Explainable AI (XAI) methods are being developed, but they are still in their early stages and face limitations in complex biological systems.

Process Complexity and Dynamics:Biopharmaceutical processes involve intricate biological interactions, non-linear relationships, and dynamic feedback loops. Static ML models trained on historical data may struggle to adapt to these complexities, leading to inaccurate predictions and potential safety concerns. Continuously learning models and reinforcement learning approaches hold promise, but their implementation requires careful design and validation in controlled settings.

Infrastructure and Integration:Biopharmaceutical companies often have legacy data systems and infrastructure that may not be readily compatible with cutting-edge ML tools. Integrating these systems, ensuring data security and compliance, and building pipelines for real-time data capture and model deployment present significant challenges. Collaboration between data scientists, engineers, and domain experts is crucial to overcome these hurdles.

Regulatory Landscape and Validation:Regulatory agencies naturally exercise caution towards incorporating novel technologies like ML into drug development and manufacturing. Establishing standardized guidelines, rigorous validation procedures, and clear documentation for ML models is essential to gain regulatory approval and ensure patient safety. Collaboration between industry, academia, and regulators is key to navigating this evolving landscape.

Ethical Considerations and Bias:Biopharmaceutical algorithms must uphold ethical principles of fairness, transparency, and accountability. Biases in data collection, model training, and interpretation can lead to discriminatory outcomes and unfair clinical trial designs. Mitigating these risks requires careful data analysis, diverse representation in research teams, and ongoing monitoring of model performance for potential biases.

Workforce Development and Training:Integrating ML into biopharmaceutical processes necessitates a skilled workforce equipped with both biological and computational expertise. Upskilling existing employees, fostering interdisciplinary collaborations, and developing educational programs tailored to this novel field are crucial for successful adoption.

Cost and Investment:Developing and implementing robust ML solutions requires significant financial investment, including infrastructure upgrades, talent acquisition, and ongoing maintenance. Balancing these costs with the potential benefits requires careful planning, cost-benefit analysis, and demonstration of clear value propositions.

Addressing these challenges necessitates a multifaceted approach:

- Data management and integration: Invest in building robust data infrastructure, standardize data formats, and prioritize data quality.
- Advanced ML techniques: Explore XAI methods, continuously learning models, and leverage domain knowledge to guide model development.
- Collaboration and partnerships: Foster collaboration between data scientists, engineers, biologists, and regulatory bodies.

- Ethical considerations: Implement measures to mitigate bias, ensure transparency, and uphold ethical principles throughout the ML lifecycle.
- Workforce development: Upskill existing employees, support interdisciplinary collaborations, and invest in educational programs.
- Cost-benefit analysis: Carefully evaluate potential benefits, optimize resource allocation, and demonstrate clear value propositions.

4)The applications of ML in early stages of monoclonal antibody development

Monoclonal antibodies (mAbs) have revolutionized healthcare, offering targeted therapies for numerous diseases. However, developing these lifesaving molecules can be a lengthy and expensive process, often taking upwards of a decade. Machine learning (ML) is emerging as a powerful tool, accelerating and optimizing various stages of mAb development, particularly in the crucial early stages.

1. Target Identification and Validation:

- Predicting Antigen-Antibody Interactions: ML algorithms can analyze vast datasets of protein structures and binding interactions to predict potential antibodies and their binding affinities to targeted antigens. This virtual screening process reduces the need for expensive and time-consuming wet-lab experiments, accelerating target identification and validation.
- Prioritizing Promising Targets: Using disease-specific data and biological pathways, ML models can prioritize target antigens with high potential for therapeutic efficacy, focusing resources on the most promising avenues.

2. Antibody Library Design and Engineering:

- Optimizing Antibody Sequences: ML algorithms can analyze existing antibody sequences with desired properties to suggest mutations that enhance affinity, specificity, and other functional characteristics. This "in silico" design accelerates the generation of high-performing antibody candidates.
- De novo Antibody Design: Deep learning models can create entirely new antibody sequences with desired properties, leveraging vast structural and functional data. This "de novo" design expands the search space beyond existing antibodies, potentially leading to novel therapeutic solutions.

3. High-Throughput Screening and Selection:

- Identifying Lead Candidates: ML models can analyze high-throughput screening data, rapidly identifying antibodies with desired characteristics like high affinity, specificity, and developability. This data-driven approach improves the selection process and reduces the risk of overlooking promising candidates.
- Predicting Off-Target Effects: ML algorithms can analyze antibody sequences and structural features to predict potential off-target binding and potential safety concerns, enabling early identification and mitigation of risks.

4. Predicting and Optimizing Downstream Processes:

- Predicting Expression and Purification Yields: ML models can analyze cell lines and expression conditions to predict antibody expression and purification yields, optimizing production processes and minimizing resource waste.
- Predicting Pharmacokinetic Properties: Using sequence and structural data, ML models can predict antibody stability, half-life, and tissue distribution, guiding formulation design and ensuring optimal in vivo performance.

5)The applications of ML in understanding, prediction, optimisation, monitoring and control of upstream processes

The biopharmaceutical industry thrives on innovation, constantly looking for ways to streamline processes, improve efficiency and deliver life-saving therapies faster.Machine learning (ML) represents a transformative force in this business sector, providing powerful tools to understand, predict, optimize, monitor and control upstream biological processes.Understanding complexity: Upstream processes, including cell culture, fermentation, and protein purification, are inherently complex and multifaceted.ML models can analyze large data sets generated from sensors, bioreactors, and analytics tools, uncovering hidden patterns and relationships that influence process performance.Clustering algorithms can identify distinct production patterns, while dimensionality reduction techniques can identify critical process parameters that influence yield and product quality.This deeper understanding enables bioprocess engineers to make informed decisions, optimize culture conditions, and proactively address potential problems.Predict the unpredictable: ML can turn uncertainty into informed predictions.By analyzing historical data and integrating real-time sensor information, models can predict process deviations, predict product quality attributes, and even estimate subsequent output .This predictive capability allows for early intervention, avoiding costly downtime and ensuring consistent product quality.For example, a recurrent neural network (RNN) can analyze bioreactor time-series data, predicting potential bottlenecks or trends in product titration, allowing proactive adjustments to maintain optimal conditions.Optimizing efficiency and quality: ML algorithms excel at finding the "sweet spot" in complex systems.Optimization algorithms, such as Bayesian optimization and reinforcement learning, can navigate the complex bioprocessing landscape, determining the optimal combination of process variables (e.g., temperature , pH, media composition) to maximize product yield, purity and overall process efficiency.This data-driven approach outperforms traditional, time-consuming testing, speeds up process development, and reduces manufacturing costs.Real-time monitoring: Continuous monitoring is critical to ensure process stability and product quality.ML algorithms can analyze real-time data streams from sensors, thereby identifying small deviations from desired parameters in real time.Anomaly detection algorithms can flag potential problems before they become severe, allowing for immediate corrective actions and avoiding costly process errors.Additionally, unsupervised learning techniques such as principal component analysis (PCA) can monitor the stability of processes, detecting even small changes that can indicate potential problems before they occur they can be manifested.Control with precision: ML-based control systems can take monitoring and forecasting insights to the next level.These systems can automatically adjust process parameters based on real-time data and predicted results, ensuring precise control and maintaining optimal conditions throughout the process.

This closed-loop control, powered by reinforcement learning, can adapt to changing conditions and disruptions in real time, thereby creating more robust and efficient biological processes .

6)The applications of ML in prediction, monitoring, control, and optimisation of downstream processes

Downstream preparing utilizes numerous refinement procedures to capture the required protein whereas dispensing with debasements related with different handle components (such as antifoam and Protein A leachate), have cell materials (counting have cell proteins and DNA), and product-related debasements (such as totals and parts) (Rathore and Kapoor, 2014). The particular unit operations utilized in downstream handling can change depending on the item being decontaminated, but regularly incorporate a combination of the taking after steps:chromatography, cleaning, viral filtration/clearance, and ultrafiltration/diafiltration. This area will audit ordinary applications of ML in different stages of a downstream handle.

Chromatography includes an imperative part to play within the filtration of biopharmaceutical items. As of now, a common filtration strategy for biopharmaceuticals includes utilizing two or more single-column chromatographic steps in progression (De Luca et al., 2020). The starting step, known as the capture step, is utilized to dispense with any non-product-related pollution, such as host-cell proteins, DNA, and lipids. Following the capture step, numerous cleaning steps are utilized to attain the required level of immaculateness for the target atom by evacuating pollutants that are product-related. These pollutions incorporate species created amid blends that have exceptionally comparable chemical characteristics to the target compound, such as truncated or deamidated species and totals. This area will cover the commonplace applications of ML in tending to the observing and expectation issues in a capture chromatography preparation. The common steps included within the capture chromatography are (Biosciences, 2001, Narayanan et al., 2021c):

Column arrangement:

The chromatography column is pressed with a gum which encompasses a tall official capacity for the target particle. The gum is ordinarily pre-treated to advance its execution and limit non-specific authority.

Test stacking:

The test is stacked onto the column, usually employing a peristaltic pump. The test is empowered to stream through the tar bed, and the target atom ties to the gum whereas other debasements pass through.

Washing:

After the test has been stacked, the column is washed with a buffer arrangement to dispense with any unbound debasements.

Elution:

The target molecule is then eluted from the column employing a particular elution buffer disturbing the interaction between the target particle and the tar. The eluted fabric is collected in a partitioned holder for advance handling.

Column recovery:

After the elution, the column is ordinarily recovered employing a cleaning arrangement to kill any remaining pollutions and prepare it for the following cycle of decontamination.Machine learning can be utilized to create prescient models that can figure the conduct of the capture chromatography handle based on authentic information, e.g., fouling columns and breakthrough profiles. This can offer assistance to

recognize potential issues some time recently they happen, permitting for proactive remedial activity. Machine learning can too be utilized to screen the capture chromatography handle in real-time, upgrading working parameters and enhancing the proficiency of the overall prepare.

Amid the capture stage, a protein blend is passed through a column containing a stationary stage (which is regularly a gum with particular ligands) outlined to specifically tie the target protein, whereas other debasements and proteins will pass through. The relationship between the concentration of a target protein (e.g., mAbs) within the column effluent and the volume of the test that has been passed through the column over time is graphically spoken to by breakthrough bends. At first, the gushing contains a moo concentration of the target protein, but as more tests are passed through the column, the concentration of the target protein in the gushing increments until it comes to a level. Breakthrough bends are valuable for evaluating the efficiency of the chromatography column, distinguishing the greatest stacking capacity, and deciding the point of breakthrough, which is when the column is not able to capture the target protein successfully. The shape and characteristics of the breakthrough bends give important information regarding the conduct of the target protein within the chromatography handle. For example, a soak breakthrough profile can show that the authoritative capacity of the chromatography gum is being exceeded, and additional steps may be required to get the required immaculateness (Sun et al., 2021). On the other hand, a more level breakthrough profile can demonstrate that the chromatography gum is underused which higher yields may be obtained with a better stacking density. Therefore, the accurate prediction of the breakthrough bends is crucial to advance different parameters of the capture chromatography preparation, such as the stacking thickness onto the column, stream rate, and elution conditions (Narayanan et al., 2021c). In expansion, by precisely foreseeing the breakthrough bends, one can reduce time and take a toll whereas expanding the abdicate and virtue of the target protein.

Field et al. (2019a) presented a half breed demonstration combining a deterministic chromatographic demonstration with online data created from Raman-based PLS gauges to estimate the mAb concentration at the outlet of a column aiming to screen the chromatographic breakthrough bends. Exploratory comes about showing that Raman-based prescient PLS models may capture the general shape of the breakthrough bends, but they come about as noisy to be utilized in commonsense applications such as prepare control. This clamor was caused by the tall levels of debasements within the collected test, coming about in covering ghostly highlights of different species such as the target monoclonal antibody, media components, host cell proteins, DNA, and tall atomic weight compounds. As a result, wide groups and vague top profiles were watched inside single spectra. On the other hand, the mechanistic show, i.eThe lumped motor demonstration (LKM), which was fittingly calibrated with an outside dataset, was able to capture the subjective shape of the breakthrough bends. However, it showed deviations from the off-line reference estimations that were as well expansive. Hence, the creators proposed an amplified Kalman channel approach, which combines the commitment of the Raman-PLS and the LKM forecasts within the last evaluated values of the channel. Within the crossover show, the LKM gives vigorous forecasts, whereas the real-time data from Raman-PLS can be utilized to update the state gauges and viably limit the LKM counterbalanced. As a result, the proposed EKF gives superior results in comparison to Raman-PLS and LKM models. In a

follow-up, Narayanan et al. (2022b) affirmed that both data-driven and robotic models are inclined to significant predispositions. The data-driven models need generalisability, whereas the unthinking models are overly compelled and may not satisfactorily speak to the fundamental marvels. To overcome these issues, the creators proposed to use a different degree of hybridisation between an LKM demonstrator and neural systems to define different cross breed models. The observational results showed that crossover models gave precise expectations of breakthroughs and inside column profiles of solid and liquid-phase concentration. This data is critical for preparation control and advancement since to advance gum use and keep up tall surrender amid the capture step, it is crucial to stack the column to its maximum capacity while avoiding any breakthrough. The exploratory results affirmed that crossover models display exceptionally small inclination and they can adjust to different frameworks with small inconstancy since of the mechanistic portion supporting strong extrapolations. Narayanan et al. (2021d) proposed a novel hybrid modeling approach for foreseeing breakthrough bends within the capture chromatography strategy, which combines neural networks and mechanistic models to memorize the highlights of the method in an unbiased way. The execution and potential of the half breed model are examined on both in-silico and exploratory datasets, and it appears that the crossover demonstrated outperformed the LKM in terms of forecast exactness and strength in extrapolating across process conditions.

Compared to upstream preparation, online or at-line checking devices are not broadly utilized in downstream handling, particularly in chromatography, and only some cases have been displayed in the writing (Rathore and Kapoor, 2014, Field et al., 2019b). The execution of online observing techniques in downstream is challenging due to exacting requirements such as tall affectability, strength, quick reaction time, tall exactness, wide energetic run, and moo constrain of discovery (LOD) with minimal recalibration needs (Roch and Mandenius, 2016). pH, conductivity, weight, mass stream, optical thickness, and single wavelength UV spectroscopy sensors are commonly utilized in chromatography. In any case, the information they give is constrained and cannot be utilized to accurately decide item concentration or other CQAs such as product-related pollutions (e.g., aggregates, fragments, isoforms) or process-related debasements (e.g., HCPs, DNA, filtered resin ligands) (Field et al., 2019b). As of now, the foremost broadly used method for quality control in the downstream handling of mAbs is tall execution fluid chromatography (HPLC) (Field et al., 2019b, Tiwari et al., 2018), which generally works in off-line mode but in few cases too at-line. Be that as it may, the HPLC has certain limitations. For occurrence, it requires either manual test dealing with or an expensive autosampler but still showing non-negligible time delays (Rüdt et al., 2017). Besides, test arrangement and examination can be time-consuming, and hence real-time decisions are not possible. In contrast, spectroscopy is an emerging method that has an extraordinary potential for in-line usage. It offers a brief estimation time extending from seconds to minutes, and it is noninvasive and nondestructive in its operation. Additionally, the spectroscopy empowers a concurrent measurement of different factors, making it a promising elective to HPLC for quality control and real-time checking in downstream processing (Rüdt et al., 2017), where quick estimation times are imperative for the chromatography steps. To take advantage of solid focuses of spectroscopy for real-time observing of downstream processes, ML calculations and multivariate information examination approaches are usually connected to extracting information from spectroscopic estimations. A basic

audit on recent patterns and a potential of spectroscopy strategies as PAT devices for biopharmaceutical downstream processing can be found in Rolinger et al. (2020).

In protein A capture chromatography, control of the stack stage is significant for proficient and particular filtration of monoclonal antibodies (Thakur et al., 2019). It plays basic parts in enhancing authoritative capacity, limiting non-specific authoritative, anticipating gum over-burdening, guaranteeing buffer compatibility, and keeping up preparation consistency. By successfully controlling the stack stage, it is conceivable to realize the ideal authoritativeness of the target protein to the column and effective filtration of high-quality monoclonal antibodies. However, the stacking stage within the protein A capture prepare is ordinarily not controlled in real-time, instep the stack volume as often as possible depends on the offline measurement of the mAb and a traditionalist column capacity computed from resin-life time considers (Rüdt et al., 2016). In this manner, Rüdt et al. (2016) proposed a unused strategy for real-time control of the stack stage and ended loading based on the online assessed values of mAb within the profluent of a Protein A capture step produced by PLS models utilizing UV/Vis assimilation spectra. To confirm the execution of the proposed strategy, the prepared PLS show was sent for real-time controlling the stack stage of a Protein A capture step through two diverse runs. The stack stage was naturally ended in case a mAb concentration in the gushing came to 50% of item breakthrough (the primary run) or 5% of item breakthrough (the moment run). The observational results appeared that in both runs, the comparing stack stages were effectively ended near to the focused on breakpoints. In another ponder, Thakur et al. (2019) created a close infrared spectroscopy stream through cells, found some time recently the channel of the stack column and alternatively at the outlet of the stack column, to capture spectra data within the gathered broth and flow-through each three seconds. These spectra are along these lines bolstered to online PLS models, which are calibrated with the reference spectra, to measure the concentration of mAb in both the gathered broth and flow-through. The proposed strategy was confirmed utilizing two exploratory setups. In setup A, as it were the mAb concentration within the stack stream was observed. The control calculation obtained concentration information each three seconds and made choices whether or not to halt the stream of the stacking pump based on the rate breakthrough assessed from the mAb concentrations some time recently in column gulf and after column outlet. In setup B, the mAb concentrations were observed in both the stack and flow-through. The control calculation was outlined to modify the valve setup and divert the flow-through fabric to the moment pass column instead of the squander tank once a certain level of mAb breakthrough was identified. The real-time stacking control strategy accomplished the ideal gum usage whereas still keeping up occasional elutions.

After the target protein is captured by the chromatography tar, a few washing and elution steps are conducted to kill pollution and get the decontaminated protein. Each chromatography step produces different divisions or eluates that incorporate different levels of the target protein and other debasements. The pooling choices are at that point performed to choose and combine divisions from distinctive chromatography runs or columns to attain the required level of immaculateness and abdicate. These choices are based on an analysis of the virtue and amount of each division, and they offer assistance to diminish batch-to-batch variability and limit the by and large handle time and fetch. Brestrich et al. (2018) illustrated an inline control procedure of pooling choices based on either the eluting protein

concentrations or the immaculateness of pools anticipated by PLS models utilizing variable pathlength UV/Vis spectra. Within the to begin with confirmation case think about where lysozyme (the target item) was filtered from cytochrome c, the pooling choice was activated when the concentration of lysozyme surpassed 2 g/L and the cytochrome c concentration fell below 1.8 g/l based on the assessed values given by the PLS models. Within the moment case, think about with respect to the division of HMWs from the mAb monomer, the pooling was activated in case the mAb monomer concentration is bigger than 2 g/l and this pooling step was ended in case the virtue fell underneath 95%. In both cases, the pool immaculateness was at that point measured using offline reference analytics to compare the viable comes about with the one anticipated by the PLS demonstration. For the to begin with case consider, a anticipated immaculateness of 99.0% was obtained by the PLS show, whereas offline analytics come about in a measured purity of 99.7%. Within the moment case ponder, the PLS model anticipated a immaculateness of 94.4%, whereas off-line analytics measured a somewhat lower immaculateness of 94.2%. The article affirmed that the proposed strategy may be conveyed for in-line control of pooling choices inside a chromatography handle.

Profundity filtration (DF) may be a as often as possible utilized strategy to clarify cell culture broth when creating mAb items due to its moo starting fetch, direct gear, ease of operation, and approval (O'Brien et al., 2012). It can work as either the beginning step (direct depth filtration) or an ensuing step within the clarification preparation. An arrangement of one or more stages of profundity channels can be conveyed, with each progressive channel dispensing with dynamically littler particles, producing a clarification arrangement (beginning with coarse and finishing with fine profundity channels) (van Reis and Zydney, 2007). In later a long time, as bioreactor processes have advanced towards longer aging times and higher cell concentrations to attain more noteworthy item titers and yields, there has been a outstanding increment in cell flotsam and jetsam and natural constituents, driving to lower cell reasonability (Agarwal et al., 2016). As a result, the stacking capacity of profundity channels is decreased. To address this, fabricating scale forms may require expansive, single-use gatherings with a channel region of a few hundred square meters. It is significant for the profundity filtration step to be steady, since any irregularity in its execution can lead to diminished productivity in downstream filtration or misfortune of the item (Noguchi et al., 2009).

A significant challenge in creating a successful profundity filtration-based clarification procedure for huge volume bioreactors is foreseeing channel stacking capacity at scale (Noguchi et al., 2009). The current approach for scaling up profundity filtration expects that the fouling component of the channel remains the same as working conditions are maintained, empowering direct scaling up of channel range from the lab to pilot to commercial scale (Agarwal et al., 2016). In any case, a channel measuring depends on a few prepared components and crude fabric properties, such as working flux variety, lot-to-lot variety within the channel, bolster inconstancy, and operation scale (Noguchi et al., 2009). To address these issues, Agarwal et al. (2016) depicted an application of counterfeit neural organization demonstrating to anticipate profundity channel stacking capacity for mAb clarification amid commercial fabricating. The proposed demonstration utilized nourish turbidity, cell check, nourish cell reasonability, flux, and time as input parameters to foresee the weight increment in a profundity filtration unit operation. The show was prepared with

exploratory information and an amazing understanding between anticipated and test comes about with a relapse coefficient of 0.98. The investigation moreover conveyed Monte-Carlo reenactments to assess the benefits of utilizing variable profundity channel measuring compared to employing a settled channel range. The experimental results showed a 10% fetched sparing in utilizing variable profundity channel measuring for different clarification parts rather than utilizing the settled channel measuring. This consideration illustrated the potential of ANN displaying to plan proficient and cost-effective manufacturing scale filtration forms.In the generation of mAb items, the viral filtration step is significant to successfully evacuate both wrapped and non-enveloped infections based on their measure. Membrane-based frameworks have ended up progressively prevalent as an exceedingly effective technology for filtration and refinement in a wide run of applications, counting the viral filtration (Virtanen et al., 2017). Billups et al. (2022) highlighted the noteworthiness of a layer pore structure and reversible mAb total arrangement on filtrate flux amid a virus removal filtration. One of the major challenges within the utilization of an infection channel film innovation is the film fouling. This fouling can compromise the infection clearance and diminish film efficiency, characterized as the item recouped per layer surface zone (Isu et al., 2022). Fouling frequently happens due to the product and process-related variations such as have cell proteins, proteases, and endotoxins rather than any rejected infection particles since the concentration of infection particles in the handle is usually several orders of size lower than that of the item (Basile and Ghasemzadeh, 2019). In addition, fouling might too affect the selectivity of the utilized film and the substance of the divisions created amid filtration, in this manner affecting the effectiveness and cost-effectiveness of the filtration handle (Virtanen et al., 2017). Hence, understanding the components of fouling is significant to controlling and limiting it (Cui et al., 2011).Recognizing and recognizing early-stage layer fouling in real-time is vital to understanding fouling marvels. Subsequently, Virtanen et al. (2017) explored the utilization of typical Raman spectroscopy as an internet device for checking membrane fouling. Such strategies capable of foulant characterisation and online observing are basic for controlling fouling in film forms. Comes about from the ponder indicated that ordinary Raman spectroscopy is a promising instrument for early-stage layer fouling observing. The strategy is both subjective and quantitative, permitting for utilizing the crest statues of the foulant to track the advance of fouling over time. Moreover, the thought illustrated that multivariate strategies, such as vital component investigation, can uncover the energetic conduct of fouling over time. The combination of typical Raman spectroscopy and PCA can give semi-quantitative results about depicting the accumulation of organic foulant on the film structure.

7) The applications of ML to problems in the product formulation and stability

The quest for stable and effective formulations lies at the heart of various industries, from pharmaceuticals and food to cosmetics and materials. Traditionally, this relied on intuition, experience, and extensive experimentation. However, the rising complexity of products and the demand for faster development times are paving the way for a new alchemist: machine learning (ML). This essay explores the burgeoning applications of ML in tackling problems related to product formulation and stability, highlighting its potential to revolutionize the field.Understanding the Formulation Landscape:Formulating a stable product involves navigating a multi-dimensional space of ingredients, their interactions, and their impact on

critical quality attributes (CQAs) like shelf life and efficacy. ML excels at pattern recognition and data analysis, making it ideal for uncovering hidden relationships within complex datasets. Techniques like dimensionality reduction can visualize high-dimensional data, revealing previously unseen connections between formulation components and their stability profiles. This understanding empowers researchers to identify promising ingredient combinations and predict potential stability issues early in the development process.Predicting Stability: A Crystal Ball for Formulators:Accurately predicting a product's shelf life under various storage conditions is a crucial yet challenging task. Traditionally, this relied on extensive stability testing, a time-consuming and resource-intensive process. ML offers a powerful alternative. By analyzing historical stability data, incorporating physicochemical properties of ingredients, and leveraging accelerated stability testing results, ML models can accurately predict shelf life under different conditions. This predictive power allows for streamlined development, optimized packaging choices, and improved inventory management.

Optimizing Formulations: A Dance of Ingredients:Finding the optimal formulation for a complex product is akin to choreographing a delicate dance of ingredients. ML algorithms, particularly optimization algorithms, can excel at this task. Given desired product characteristics and stability constraints, these algorithms can efficiently explore vast formulation spaces, suggesting optimal combinations that maximize stability while considering factors like cost and manufacturability. This capability significantly reduces the time and resources required for traditional optimization approaches, accelerating product development and reducing waste.Monitoring Stability in Real-Time: A Vigilant Eye:Maintaining product stability throughout its lifecycle is crucial for ensuring safety and efficacy. ML can be employed for real-time monitoring of stability during storage and distribution. For instance, sensor data collected from smart packaging materials, coupled with ML-powered anomaly detection algorithms, can identify subtle changes in product characteristics that could indicate potential degradation. This early warning system allows for timely interventions, minimizing product recalls and ensuring consumer safety.

8)Opportunities for potential applications of ML towards BioPharma 4.0

A digital twin (DT) is a virtual representation of a physical system that mimics its behavior and dynamics. In biopharmaceuticals, DTs aim to create a digital technology that can access all available information in process data archives across different scales and sites, enabling real-time information exchange with the complete process control system and all human stakeholders. DTs can transform data independently of format, capture complexity, and simulate various future scenarios to derive relevant decisions. They can improve process efficiency, prediction, decision making, reduce costs, risk analysis, and increase product quality.Park et al. (2021) introduced a bioprocess digital twin platform that integrates physical and virtual systems through a data management system. The bioprocess DT includes physical operations of cell culture, real-time bioreactor monitoring, data processing, and management, as well as mechanistic and ML modeling of both cells and bioprocesses. The combination of mechanistic and data-driven modeling approaches can effectively identify bottlenecks in the process and suggest operational strategies to manage product quality. However, the integration of virtual and physical plants in biopharmaceutical manufacturing has not yet been fully developed.To develop DT solutions, it is

essential to deploy PAT tools to measure multiple process parameters and quality attributes in the bioreactor and collect a large set of cell culture data during the monitoring process of the physical systems.The necessity for a huge sum of preparing information could be a crucial impediment of machine learning (Mowbray et al., 2021). Be that as it may, within the biopharmaceutical industry, datasets regularly have the next number of estimations than perceptions (Severson et al., 2015). Regularly, estimations are taken as it were amid a number of arranged times as the group moves through the generation prepare, and few imitates are conducted due to time and taken a toll restrictions. This will posture critical challenges when applying machine learning to little datasets with occasional criticism in bioprocess improvement and fabricating information. To construct tall performing prescient models, it is required to plan particular calculations for little information sets. Regularization procedures have been perceived as potential arrangements for such issues, as they can handle both input determination and show estimation at the same time to dodge overfitting (Pampuri et al., 2011). Severson et al. (2015) combined a regularization strategy, i.e., the versatile net, with Monte Carlo examining procedures to construct an compelling learning show from restricted preparation datasets. The flexible net with Monte Carlo testing calculation blends the points of interest of the versatile net calculation, which empowers synchronous show choice and parameter estimation, with the quality of Monte Carlo inspecting to avoid potential information overfitting, coming about in a dependable, reasonable, and direct prepare model.The moment strategy to bargain with the restricted sum of training data is the utilize of different data augmentation methodologies.

9)Conclusions

The biopharma industry is experiencing change due to the requirement for moved forward medications and more proficient improvement and fabricating forms. This is often persuaded by the improvement of progressed innovations such as computational control, keen sensor advances, information administration and communication frameworks, and expository apparatuses. A strategic initiative for advanced change is essential to require full potential advantage of these innovations. This requires the execution of ML-based models for information investigation and expectations, beside robotization and IoT for framework networks (Puranik et al., 2022). ML-based models have the potential to upgrade operational methods, prepare efficiency, and progress efficiency in less time, empowering superior decision-making amid bio-manufacturing. Joining ML and AI-based foundations deliberately offers the potential to convert the bioprocesses for biopharmaceutical advancement.

This paper displayed a careful evaluation of how ML has utilized totally different stages of the improvement and fabricating forms for mAbs and helpful antibodies. Furthermore, it looks to distinguish the troubles related to bioprocesses and prepare information and to propose potential inquiries about roads in this space towards building an advanced change stage for the biopharmaceutical industry. A characteristic of bioprocess information is that it continuously includes information that underlies the information, instead of the information itself, upon which ML models are built. In this manner, a basic note with respect to the applications of ML within the biopharmaceutical field is that notwithstanding of which machine learning calculation is utilized for bioprocess information classification or relapse, their results will not be valuable in biopharmaceutical inquire about in the event that they cannot offer extra

biophysical bits of knowledge or contribute to space information era and understanding. As a result, basically depending on machine learning to reveal the covered up information of bioprocess information may be excessively idealistic, given the data-driven nature of this approach. Subsequently, the foremost basic step towards empowering the persistent application of machine learning in future considerations is to progress the accuracy of forecasts and decrease vulnerability by coordination bioprocess information into the show building preparation utilizing distinctive strategies. In addition to exploring other advanced machine learning methods, it is too imperative to provide the need to develop novel robotic models for biopharmaceutical forms in both upstream and downstream preparation. By and large, it is accepted that the developing connections between data-driven, robotic, and crossover demonstrating will proceed to be productive and complementary to each other. The shrewd integration of these demonstrating strategies will also allow for the improvement of advanced twins, which can essentially speed up the improvement of biopharma from distinguishing lab-scale plans of tests and upgrade, screen, and control industrial-scale operations.

REFERENCES

1. Abu-Absi N.R., Kenty B.M., Cuellar M.E., Borys M.C., Sakhamuri S., Strachan D.J., Hausladen M.C., Li Z.J.Real time monitoring of multiple parameters in mammalian cell culture bioreactors using an in-line Raman spectroscopy probe,Biotechnol. Bioeng., 108 (5) (2010), pp. 1215-1221, 10.1002/bit.23023
2. Adam S.P., Alexandropoulos S.-A.N., Pardalos P.M., Vrahatis M.N., "No free lunch theorem: A review",Approx. Optim. Algorithms Complexity Appl., 145 (2019), pp. 57-82
3. Agarwal H., Rathore A.S., Hadpe S.R., Alva S.J,".Artificial neural network (ANN)-based prediction of depth filter loading capacity for filter sizing",Biotechnol. Prog., 32 (6) (2016), pp. 1436-1443
4. Akbar R., Bashour H., Rawat P., Robert P.A., Smorodina E., Cotet T.-S., Flem-Karlsen K., Frank R., Mehta B.B., Vu M.H., Zengin T., Gutierrez-Marcos J., Lund-Johansen F., Andersen J.T., Greiff V. "Progress and challenges for the machine learning-based design of fit-for-purpose monoclonal antibodies",mAbs, 14 (1) (2022), pp. e2008790:1-35
5. T. T. Khuat, R. A. Bassett, E. Otte, A. Grevis-James, and B. Gabrys, "Applications of machine learning in antibody discovery, process development, manufacturing and formulation: Current trends, challenges, and opportunities," Mar. 01, 2024. https://doi.org/10.1016/j.compchemeng.2024.108585
6. Akbar R., Robert P.A., Pavlović M., Jeliazkov J.R., Snapkov I., Slabodkin A., Weber C.R., Scheffer L., Miho E., Haff I.H.k., *et al.,"*A compact vocabulary of paratope-epitope interactions enables predictability of antibody-antigen binding",Cell Reports, 34 (11) (2021), Article 108856
7. Aledo J.C., Canton F.R., Veredas F.J.,"A machine learning approach for predicting methionine oxidation sites",BMC Bioinformatics, 18 (1) (2017), pp. 1-14
8. Algorri M., Abernathy M.J., Cauchon N.S., Christian T.R., Lamm C.F., Moore C.M.,"Re-envisioning pharmaceutical manufacturing: increasing agility for global patient access",J. Pharm. Sci., 111 (3) (2022), pp. 593-607

9. Ali A.R., Budka M., Gabryś B.,"Towards meta-learning of deep architectures for efficient domain adaptation",PRICAI 2019: Trends in Artificial Intelligence: 16th Pacific Rim International Conference on Artificial Intelligence, Cuvu, Yanuca Island, Fiji, August 26–30, 2019, Proceedings, Part II 16, Springer (2019), pp. 66-79
10. Alwis D.M.D., Dutton R.L., Scharer J., Moo-Young M.,"Statistical methods in media optimization for batch and fed-batch animal cell culture,Bioprocess Biosyst. Eng., 30 (2) (2007), pp. 107-113,doi:10.1007/s00449-006-0107-7
11. Azubuike C.C., Edwards M.G., Gatehouse A.M., Howard T.P.,"Applying statistical design of experiments to understanding the effect of growth medium components on Cupriavidus necator H16 growth",Appl. Environ. Microbiol., 86 (17) (2020), pp. e00705-20
12. S. Badr and H. Sugiyama, "A PSE perspective for the efficient production of monoclonal antibodies: integration of process, cell, and product design aspects," *Current Opinion in Chemical Engineering*, vol. 27, pp. 121–128, Mar. 2020, doi: 10.1016/j.coche.2020.01.003.
13. M. Bailly *et al.*, "Predicting Antibody Developability Profiles Through Early Stage Discovery Screening," *mAbs*, vol. 12, no. 1, Jan. 2020, doi: 10.1080/19420862.2020.1743053.
14. Maglogiannis, J. Macintyre, and L. Iliadis, *Artificial Intelligence Applications and Innovations*. Springer, 2022. [Online]. Available: http://books.google.ie/books?id=d_AszwEACAAJ&dq=Bakirov+R.,+Fay+D.,+Gabrys+B.%0AAutomated+adaptation+strategies+for+stream+learning%0AMach.+Learn.,+110+(6)+(2021),+pp.+1429-1462&hl=&cd=3&source=gbs_api