AI-Driven Pharmaceutical Manufacturing: Leveraging Datarobot and GenAI / Agentic AI For Predictive Modeling and Process Optimization

Nivedha Sampath

Platform Engineer at Takeda Pharmaceuticals, Boston, USA

Abstract: ***Objectives**: To analyze the application of artificial intelligence technologies-DataRobot, Generative AI, and Agentic AI-for predictive modeling and process optimization in pharmaceutical manufacturing, with a focus on enhancing quality stability and reducing costs. **Study Design**: A comprehensive applied study involving the systematization of practical case studies and an analytical comparison of manufacturing solutions within the context of the industry's transition to the "Pharma 4.0" model. **Setting and Duration**: The work is based on a synthesis of industrial and academic sources covering the period from 2022 to 2025, focused on the implementation of digital factories and intelligent platforms. **Methodology**: The methodology included a review of publications, thematic coding of AI application examples, reconstruction of production chains, and classification of tasks by cycle stages (biotechnological processes, quality control, logistics). **Results**: The study systematized areas of predictive analytics, identified use cases for generative models for synthetic data and formulations, and described the architecture of agentic systems for monitoring, Corrective and Preventive Action (CAPA), and resource planning. The findings confirm accelerated technology development, improved process stability, and reductions in product defects and unplanned downtime. **Conclusion**: The combined use of predictive, generative, and agentic solutions forms the foundation of the digital factories of the future and ensures a sustainable increase in the efficiency of pharmaceutical manufacturing.*

Keywords: pharmaceutical manufacturing, DataRobot, predictive modeling, Generative AI, Agentic AI, Quality by Design, digital factory, quality control, synthetic data, manufacturing analytics

1. Introduction

In modern pharmaceutical manufacturing, artificial intelligence (AI) technologies are increasingly being applied to enhance the efficiency of drug development and production. The last five years have seen particularly intensive development in these approaches, driven by the need to accelerate the market entry of new drugs, reduce costs, and improve product quality control. The relevance of this topic is dictated by the global pharmaceutical industry's transition to the "Industry 4.0" concept, which involves the digitalization and intellectualization of production processes (DataRobot, 2023a). The objective of this work is to analyze the application of the DataRobot platform and modern Generative AI and Agentic AI technologies in the pharmaceutical sector for predictive modeling and the optimization of manufacturing processes. In line with this objective, the following tasks are addressed:

- To characterize the capabilities of DataRobot in the context of pharmaceutical manufacturing,
- To describe examples of the use of generative and agentic AI for modeling and managing processes,
- To evaluate the impact of these technologies on the efficiency and quality of pharmaceutical production.

2. Methods and Materials

The source corpus was formed to compare industrial AI practices with the tasks of pharmaceutical technology and compliance. DataRobot provided descriptions of AutoML, simulation, and predictive scheduling for manufacturing scenarios (DataRobot, 2023a); DataRobot offered an overview of product modules and deployment options on the shop floor (DataRobot, 2023b); and DataRobot detailed the infrastructure principles for agentic applications and action orchestration (DataRobot, 2025). M. Boskabadi presented the concept of industrial agentic AI and its connection to generative modeling in complex systems, which was used as the framework for the agent architecture (Boskabadi et al., 2025). N. Hussain reviewed predictive approaches in logistics and market analytics, which were applied in describing inventory planning and production schedules (Hussain et al., 2024). D.R. Serrano summarized the use of AI in drug discovery and drug delivery, which allowed for the transfer of PAT, QbD, and neural network parameter selection practices to granulation, drying, and pressing operations (Serrano et al., 2024). D. Staunton described examples of agentic AI implementation in life-science manufacturing, which formed the basis for scenarios in laboratory automation and CAPA processes (Staunton, 2025). L.K. Vora characterized the engineering applications of AI in pharmaceutical technologies and dosage form design, which was used in the systematization of tasks for visual inspection and stabilization of batch-to-batch variability (Vora et al., 2023).

The methodology included content analysis of documents, thematic coding of case studies, a comparative analysis of procedures, mapping of tasks to cycle stages, analytical synthesis, verification of conclusion consistency across sources, and reconstruction of data flows for production loops. The final step was the assembly of three systematizations and a protocol for the reproducible transfer of practices into a production environment.

3. Results and Discussion

One of the key areas for AI applications is the prediction of various parameters and outcomes of the manufacturing process. The DataRobot platform, a leader in the field of

AutoML (automated machine learning), provides tools for the rapid creation and deployment of predictive models based on production data. For example, with DataRobot, enterprises can develop models to predict the yield of a biotechnological process, optimize the composition of raw materials, or predict the probability of defects in a drug batch (DataRobot, 2023b). According to the company, the use of such models allows for maximizing product yield and reducing costs through the optimal selection of parameters—the so-called profitability and yield matrix, where the ideal combination of components and process conditions is determined (DataRobot, 2023b). Furthermore, the integration of prediction algorithms into production lines helps reduce unplanned equipment downtime: for instance, analyzing sensor data and equipment operation logs makes it possible to predict failures and perform maintenance preventively. As a result, one enterprise that implemented AI models for failure prediction achieved a significant reduction in downtime and associated losses. Applying machine learning methods to process data also enables advanced process control—dynamically adjusting conditions in real time. According to a review by FDA experts, combining real sensors with predictive models in a process control loop makes it possible to predict the course of a technological operation and maintain optimal conditions, which increases product quality stability (DataRobot, 2023a). In general, machine learning algorithms are now successfully used to model complex non-linear relationships between production parameters and drug quality, which is difficult to achieve with traditional methods. As noted in a recent study, the implementation of AI accelerates the development and scaling of technologies, allowing for the faster discovery of optimal formulations and regimes at the process design and scale-up stages (DataRobot, 2023a). Thus, AI-based predictive modeling has become a powerful tool for Quality by Design in pharmaceuticals—an approach where quality is built in by process design rather than being controlled only at the output (Vora et al., 2023). A systematization of these approaches is presented below (Table 1).

Table 1: Classification of Predictive Modeling Tasks in Pharmaceutical Manufacturing: Objectives, Data, Algorithms, Solutions (compiled by the author based on (DataRobot, 2023a; Serrano et al., 2024; Vora et al., 2023)

Production Cycle Stage	Prediction/Optimization Objective	Data Sources	Typical Algorithms	Control Loop Solutions
Biotechnological Reactor	Predict yield and quality stability	Process sensors, equipment logs, and laboratory protocols	AutoML ensembles, gradient boosting, and regularized regressions	Proactive parameter correction, QbD (Quality by Design) experiment planning
Granulation/ Drying/Pressing	Stabilize dosage form characteristics	PAT (Process Analytical Technology) streams, moisture and temperature profiles, QC (Quality Control) reports	Neural network models combined with response surface methodology	Adjustment of operational modes, reduction of batch-to-batch variability
Unit-Dose Quality Control	Detect defects and deviations	Image and video streams from lines, defect standards	Computer vision, convolutional networks, anomaly detectors	Automated rejection of non-conforming units, routing for re-inspection
Maintenance	Prevent equipment failures	Telemetry, vibration, and acoustic signals, PLC (Programmable Logic Controller) logs	Remaining useful life prediction, probabilistic failure models	Scheduling maintenance before stoppage, reducing unplanned downtime
Supply Chain	Align production with the demand forecast	Sales, logistics schedules, and external factors	Time-series models, gradient-based methods	Batch and inventory planning, reducing the risk of shortages

AI technologies demonstrate high potential in optimizing both individual technological stages and the entire drug supply chain. Through deep learning and other algorithms, large volumes of trial and monitoring data can be analyzed to identify hidden factors affecting the quality of the dosage form. For example, in the production of solid tablet forms, neural networks and evolutionary algorithms are used to select optimal conditions for granulation, drying, and pressing, ensuring the uniformity of product characteristics (Vora et al., 2023). Modern research shows that using neural network models in conjunction with statistical methods (response surface methodology, etc.) allows for a better understanding of the influence of each parameter on the critical quality attributes of a drug and for finding parameter combinations that reduce variability between batches (Hussain et al., 2024; Vora et al., 2023). This is especially important for biotechnological drugs and complex dosage forms, where the traditional empirical approach does not always successfully prevent defects. The application of AI leads to increased process stability: for instance, one review notes that algorithms, by analyzing data in real time, can detect trends in quality changes and suggest corrective actions even before parameters go out of tolerance (Vora et al., 2023). As a result, a reduction in the proportion of non-conforming products is observed, and drug recalls from the market occur less frequently.

Furthermore, AI plays a crucial role in the manufacturing analytics system—for example, computer vision is used for the automated quality control of tablets and capsules. Machine vision systems, trained on samples of defects, can inspect each unit of production at high speed for cracks, chips, color changes, and other deviations. Automating quality control with AI not only speeds up this process but also provides a more objective assessment compared to a human inspector, minimizing the risk of a defect being missed. In parallel, AI contributes to the optimization of areas adjacent to production, such as logistics and inventory management. Predictive models integrated into supply chain management platforms forecast the demand for drugs and allow for the optimal planning of production and delivery volumes. Consequently, storage costs and the risk of shortages of important medicines in pharmacies are reduced (Serrano et al., 2024). In sum, data from recent years indicate that the application of AI technologies leads to increased efficiency and reliability in pharmaceutical manufacturing and facilitates the industry's transition to the "Pharma 4.0" concept—the integration of automation, data, and AI to create smart factories.

Alongside predictive modeling, generative models and agentic AI applications have garnered increasing interest in recent years. Generative AI can create new data or variants based on trained samples. In the context of drug manufacturing, this finds application, for example, in generating optimized molecular structures of drug candidates, as well as in synthesizing new formulations and experimental conditions. Although the primary field for generative AI is the early stages of drug discovery, there are already examples of its integration into the process development stage. For instance, generative models can be used to simulate various process scenarios, creating synthetic data to train predictive algorithms when real experimental data is scarce. Additionally, large language models (LLMs) are beginning to be used for analyzing scientific texts and reports, automatically generating production progress reports, and even supporting line operators with interactive guidance. DataRobot, as a universal AI platform, has already integrated generative AI capabilities alongside classic machine learning. The range of applications is summarized in a thematic matrix (Table 2).

Table 2: Generative AI and Synthetic Data in Technology Cycle Stages: Objectives, Methods, Artifacts, Application (compiled by the author based on (DataRobot, 2025; Serrano et al., 2024; Vora et al., 2023)

Stage	Primary Objective	Generative Methods	Resulting Artifacts	Application in Manufacturing
Formulation and Process Design	Select combinations of parameters and conditions	Conditional generators, Bayesian optimization, evolutionary approaches	Sets of alternative operational profiles and formulations	Reducing the number of physical experiments, accelerating stabilization
Model Validation and Robustness	Fill data gaps and address rare scenarios	Synthetic datasets, digital simulators, and perturbation generation	Scenarios of deviations and rare events	Stress-testing quality control and management loops
Documentation and Knowledge	Automate reports and procedures	LLMs, retrieval-augmented generation	Draft reports, SOP templates, excerpts from registries	Accelerating compliance and audits
Operator Support	Navigate protocols and provide guidance	Conversational LLM agents	Interactive instructions, checklists	Reducing operator errors on the line
Production and Supply Planning	Align schedules and risks	Generation of demand and supply scenarios	Alternative plans and priorities	More flexible rescheduling under resource constraints

The concept of Agentic AI, which has developed between 2023 and 2025, deserves special attention. Unlike individual models, agentic AI involves a set of intelligent agents capable of autonomously performing sequences of tasks by making decisions based on data and objectives. The DataRobot environment now includes an infrastructure for developing agentic AI applications, allowing for the combination of various models (both generative and predictive) and tools into a single pipeline (Staunton, 2025). In pharmaceutical manufacturing, the agentic approach opens new possibilities for automating complex workflows. For example, in laboratory practice, agentic AI can monitor instrument readings and automatically plan interventions: one agent continuously tracks reactor sensor data, another compares it with a predictive model and decides whether to adjust the temperature or pH, and a third, upon detecting an anomaly, independently requests human confirmation to stop the process. As experts note, agentic AI can reduce equipment downtime and increase laboratory throughput through predictive maintenance—agents anticipate failures and schedule repairs at the optimal time (Boskabadi et al., 2025; Staunton, 2025).

Furthermore, agentic systems successfully handle the task of intelligent planning: in situations where multiple production lines compete for resources (raw materials, personnel), AI agents can dynamically reallocate resources and reconfigure shift schedules to ensure no unit is idle and no bottlenecks occur. For example, upon implementing Agentic AI at a pharmaceutical plant, agents monitored warehouse inventory levels, logistics schedules, and shop floor load, and upon detecting a risk of raw material delivery delay, they automatically reprioritized production to other products, thus avoiding a general production stoppage (Staunton, 2025). Similarly, upon detecting a slowdown in the laboratory stage (e.g., delays in analytical testing), agentic AI can temporarily reallocate personnel or equipment to smooth out the delay. An important area is the integration of Agentic AI with quality control systems: agents can not only detect deviations but also independently initiate investigations into the causes of defects, requesting necessary data from various enterprise information systems and preparing a draft report for the quality department. Such agentic control ensures continuous improvement—the system learns from each incident and proposes measures to prevent similar deviations in the future. The structure of an agentic system is summarized in an overview profile (Table 3).

Table 3: Architectural Elements of Agentic AI in Pharmaceutical Manufacturing: Agent Classes, Responsibilities, Inputs, and Integration Points (compiled by the author based on Boskabadi et al., 2025; DataRobot, 2025; Staunton, 2025)

Agent Class	Responsibility	Inputs/Knowledge	Actions/Effects	Integration Points
Process Monitoring	Continuously track metrics and compare against predictions	PAT/SCADA streams, predictive models, control charts	Notifications, corrective recommendations, scenario initiation	DCS/MES, operator interfaces
QC and Investigations	Identify root causes of non-conformance, prepare CAPA	LIMS, reports, equipment logs, specifications	Draft CAPAs, task routing, and evidence collection	eQMS, LIMS, DMS
Planning and Resources	Synchronize schedules, shifts, and raw materials	ERP, WMS, demand forecasts, line statuses	Reschedule production, prioritize batches	ERP/MES, warehouse systems
Maintenance	Predictive reliability support	Telemetry, failure maps, service history	Create work orders, schedule maintenance windows	CMMS, calendar systems
Compliance	Prepare regulatory reports	SOPs, parameter registries, deviation logs	Collect and format the evidentiary basis	DMS, eQMS, reporting portals
Knowledge and Operator Assist	Answer questions, summarize data	Knowledge bases, formulations, tech notes	Step-by-step instructions, conversational guidance	HMI, chat interfaces, knowledge portal

Despite its novelty, Agentic AI technology is already being considered the next stage in the development of digital pharmaceutical manufacturing. It builds on the solid foundation of predictive models and extends them, allowing not only for prediction but also for autonomous action based on those predictions. The DataRobot platform offers enterprises a kind of sandbox for developing agentic applications with an open architecture, support for connecting any models and tools via API, and a built-in data store and vector databases for knowledge management (DataRobot, 2025). This facilitates the creation of specialized agents for the needs of a specific production facility. For example, a pharmaceutical company can implement an agent responsible for regulatory compliance: such an agent would analyze process data and generate reports for regulators, verifying that all parameters are within limits and that deviations are properly documented. Another agent could handle personalization—for instance, generating production reports for specific batches for important clients or adapting the production plan to a sales forecast generated by a marketing model.

The results of the analysis show that the use of AI, particularly DataRobot solutions and GenAI/Agentic AI technologies, leads to significant changes in all aspects of drug manufacturing. Predictive modeling based on machine learning has already become firmly established in the practice of large pharmaceutical companies, allowing them to shift from reactive management (correcting problems after they occur) to a proactive style—preventing deviations and optimizing parameters before problems arise. This aligns with modern scientific views: a recent review notes that AI "catalyzes a profound transformation of the pharmaceutical industry" across all stages of a drug's lifecycle, from discovery to post-marketing surveillance. In fact, over a short period, there has been a transition from isolated experiments to the systemic integration of AI into production processes. At the same time, academic studies confirm the practical effect: according to a 2023 review, the implementation of machine learning algorithms has reduced technology development time and increased the efficiency of quality control in several pharmaceutical manufacturing facilities.

However, the limitations of current AI solutions must also be critically assessed. First, the quality of predictions and autonomous actions depends entirely on the data on which the models are trained. Pharmaceutical processes are characterized by high complexity and variability of raw materials (especially in biotechnology), so reliable training requires extensive and representative data. In practice, collecting such data is difficult—processes are validated, and frequent experiments outside narrow operating ranges are not permitted. This creates the risk that a model will inaccurately predict process behavior outside of familiar conditions. Second, there are regulatory aspects: the pharmaceutical industry is strictly regulated, and the use of AI for making critical decisions (e.g., an agent independently correcting a process) raises questions of GMP compliance. Regulators, including the FDA, are still approaching this cautiously—discussion documents note that methodologies for validating AI models and ensuring their transparency need to be developed. Thus, before agentic AI can fully and autonomously manage a pharmaceutical plant, the problem of trust in such systems, their explainability, and their verification across all possible scenarios must be solved.

Nevertheless, it is already clear that humans and AI can effectively complement each other. A hybrid approach seems optimal: AI serves as a "smart assistant," proposing solutions and handling routine tasks, while the final critical decision remains with the specialist. For example, DataRobot provides a "Model Insights" functionality—model explanations and visualizations of feature importance, which allows engineers to understand why an algorithm recommends a particular action. This increases the acceptance of AI in the conservative environment of process engineers. Similarly, agentic systems can operate in a semi-automatic mode: an agent makes a decision but executes it only after operator confirmation. This approach is currently being implemented on pilot lines, and personnel feedback is positive—people see agents as a help, not a threat to their expertise.

On the one hand, investments are required in data infrastructure, personnel training, and licenses for platforms like DataRobot. On the other hand, the long-term return is expressed not only financially (reduction in defects, faster product launch) but also in an increase in the company's scientific and technical potential. An organization that has accumulated data and trained models acquires know-how that is difficult for competitors to copy. In this sense, the countries and companies that are the first to master AI in pharmaceuticals gain a strategic advantage. It is no coincidence that the FDA is actively stimulating discussion about the use of AI in manufacturing, seeking to adapt the regulatory framework to innovation.

Finally, it is important to note that DataRobot is not the only platform, and Generative/Agentic AI is not a panacea. Competing solutions (e.g., C3.ai, IBM Watson, proprietary developments by large pharmaceutical companies) are also showing success. The general trend is that the industry is moving towards greater automation of knowledge: accumulated scientific data, formulas, and formulations are being combined with production data in unified data lakes, on top of which AI algorithms operate. In the coming years, we can expect the emergence of fully digital "shadows" of pharmaceutical plants—digital twins that allow for the simulation of experimental scenarios and process optimization in a virtual environment before they are implemented on real equipment. Elements of this are already visible in the application of generative models: for example, virtual sensors are being created that generate plausible data to test the robustness of a quality control system. Agentic AI, being essentially an orchestrator of multiple models, can play a key role in managing these digital factories of the future.

4. Conclusion

The development of artificial intelligence methods is fundamentally changing the approaches to drug manufacturing. The conducted research has shown that the DataRobot platform and similar solutions enable pharmaceutical companies to effectively use predictive algorithms for modeling and optimizing processes, from the stages of technological development to the operational management of production. Generative models and agentic AI applications, which have emerged on the wave of recent advancements (2022–2025), complement these capabilities by introducing an element of autonomy and "intelligence" into production systems. The main conclusions are as follows: (1) predictive modeling with AI significantly improves quality stability and productivity in pharmaceutical manufacturing through proactive parameter control and timely equipment maintenance; (2) the integration of generative approaches opens new avenues for optimization (e.g., process simulation, automated processing of text and knowledge), which accelerates the rollout of new technologies and improves documentation; (3) agentic AI has the potential to become the connecting link between disparate models and systems, providing holistic, autonomous management of multi-component processes—in practice, this translates to increased production flexibility, reduced response time to off-spec situations, and better resource utilization.

The scientific significance of the obtained results lies in demonstrating the effective combination of various AI technologies (AutoML, deep learning, and multi-agent systems) to solve complex problems in pharmaceutical technology. The practical significance is confirmed by initial cases of successful implementation: companies report reductions in production cycles, lower defect rates, and increased process yields after integrating AI solutions into critical control loops. Russian and international enterprises in the pharmaceutical sector, striving for global competitiveness, have an objective interest in actively mastering these methods. Summarizing the results, it can be concluded that the application of DataRobot, Generative AI, and Agentic AI in drug manufacturing has ceased to be an experiment and has become a working tool for the process engineer. Ahead lies the further deepening of this trend: the emergence of fully digital factories where AI continuously learns and improves the process. The transition of the pharmaceutical industry to these new "rails" of intelligent technologies promises not only economic benefits but also direct societal benefits, as the faster and higher-quality production of medicines means better access to modern therapy for patients worldwide.

Refereces

[1] Boskabadi, M., Cao, Y., Khadem, B., Clements, W., Gerek, Z., Reuthe, E., Sivaram, A., Savoie, C., & Mansouri, S. (2025). Industrial agentic AI and generative modeling in complex systems. *Current Opinion in Chemical Engineering, 48*, 101150. https://doi.org/10.1016/j.coche.2025.101150

[2] DataRobot. (2023). *AI for manufacturing – Streamline production workflows with AI-driven simulation and predictive scheduling* [PDF]. U.S. Food and Drug Administration. https://www.fda.gov/media/165743/download

[3] DataRobot. (2023). *Manufacturing solutions*. https://www.datarobot.com/solutions/manufacturing

[4] DataRobot. (2025). *Agentic AI product page*. https://www.datarobot.com/product/agentic-ai

[5] Hussain, N., Austin, B., Adepoju, P., & Afolabi, A. (2024). AI and predictive modeling for pharmaceutical supply chain optimization and market analysis. *International Journal Of Engineering Research And Development*.

[6] Serrano, D. R., Luciano, F. C., Anaya, B. J., Ongoren, B., Kara, A., Molina, G., Ramirez, B. I., Sánchez-Guirales, S. A., Simon, J. A., Tomietto, G., Rapti, C., Ruiz, H. K., Rawat, S., Kumar, D., & Lalatsa, A. (2024). Artificial intelligence (AI) applications in drug discovery and drug delivery: Revolutionizing personalized medicine. *Pharmaceutics, 16*(10), 1328. https://doi.org/10.3390/pharmaceutics16101328

[7] Staunton, D. (2025). Agentic AI's role in life science manufacturing. *Pharmaceutical Manufacturer Media*. https://pharmaceuticalmanufacturer.media/pharmaceutical-industry-insights/digital-health-in-pharma/agentic-ais-role-in-life-science-manufacturing

[8] Vora, L. K., Gholap, A. D., Jetha, K., Thakur, R. R. S., Solanki, H. K., & Chavda, V. P. (2023). Artificial intelligence in pharmaceutical technology and drug delivery design. *Pharmaceutics, 15*(7), 1916. https://doi.org/10.3390/pharmaceutics15071916