Journal of Intelligent Systems Engineering and Applied Data Science
ISSN 2xxx-xxxx
Vol. 1., No. 1, 2024, pp. 8-14

A review of Naive Bayes and decision tree methods for predicting particle size distribution in pharmaceutical manufacturing

Tole Sutikno[1*], Abdul Fadlil[2], Sunardi[3], Imam Riadi[4], Anton Yudhana[5]
Master Program of Electrical Engineering, Fac. of Industrial Tech., Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[1]tole@te.uad.ac.id*, [2]fadlil@mti.uad.ac.id, [3]sunardi@mti.uad.ac.id, [4]imam.riadi@mti.uad.ac.id, [5]eyudhana@ee.uad.ac.id
[*]*Corresponding author*

Abstract

Pharmaceutical manufacturing relies heavily on accurate particle size distribution prediction for drug efficacy, bioavailability, and patient safety. Machine learning algorithms like Naïve Bayes and Decision Tree have gained popularity for their ability to forecast complex data patterns and make informed predictions. However, Naïve Bayes assumes all features are independent, which may compromise the accuracy of predictions in certain scenarios. Researchers have explored hybrid approaches that combine Naïve Bayes with other machine learning algorithms, such as decision trees. The Decision Tree method, which is based on strong data mining methods like multivariate data analysis (MVDA), could help predict important quality factors like particle size distribution. By integrating innovative technologies like nanoelectrodes, the Decision Tree method can enhance efficiency and precision in predicting particle size distribution within pharmaceutical formulations. Accurate particle size distribution prediction is crucial for ensuring the quality and efficacy of pharmaceutical products. Future research should focus on combining Naïve Bayes and Decision Tree methods with advanced machine learning techniques, focusing on feature selection techniques and real-time monitoring and control systems within pharmaceutical manufacturing processes.

Keywords: machine learning, Naïve Bayes, decision tree, particle size distribution, pharmaceutical manufacturing

1. INTRODUCTION

Pharmaceutical manufacturing research has made significant strides in understanding and optimizing the particle size distribution of drug compounds. This crucial parameter plays a pivotal role in drug efficacy, bioavailability, and patient safety. As a result, accurately predicting particle size distribution is critical for developing high-quality pharmaceutical products. In recent years, machine learning algorithms such as Naïve Bayes and Decision Tree have gained popularity for their ability to forecast complex data patterns and make informed predictions [1]–[17]. By leveraging these methods, researchers have been able to improve the accuracy and efficiency of particle size distribution prediction models [17]. This review aims to critically evaluate the effectiveness of Naïve Bayes and Decision Tree methods in predicting particle size distribution in pharmaceutical manufacturing, shedding light on their strengths, limitations, and potential applications in the industry.

The pharmaceutical industry places significant emphasis on controlling the particle size distribution of active pharmaceutical ingredients (APIs) and excipients during manufacturing processes [18]–[20]. This is crucial, as particle size can influence the drug's dissolution rate, stability, bioavailability, and ultimately, its therapeutic efficacy. Understanding the background behind particle size distribution in pharmaceutical manufacturing is essential for ensuring product quality and consistency. Studies have shown that particle size distribution deviations can lead to variations in drug performance and may impact patient safety [17]. Therefore, pharmaceutical companies invest substantial resources in research and development to optimize particle size distribution for each drug formulation. Moreover, regulatory bodies such as the Food and Drug Administration (FDA) have strict guidelines in place to ensure that pharmaceutical products meet quality standards, including specifications related to particle size distribution [21].

Understanding and predicting particle size distribution in pharmaceutical manufacturing is crucial for ensuring product quality and efficacy [17]. Accurate measurement and control of particle size distribution can impact the bioavailability, stability, and effectiveness of a drug formulation. By predicting particle size distribution, pharmaceutical companies can optimize their manufacturing processes, reduce costs, and enhance product performance. Additionally, predicting particle size distribution enables researchers to design more effective drug delivery systems and formulations tailored for specific medical conditions. Naïve Bayes and Decision Tree methods have emerged as powerful tools for predicting particle size distribution in pharmaceutical manufacturing due to their ability to handle complex and high-dimensional data sets effectively. Through the application of these machine learning techniques, researchers can make informed decisions regarding formulation development, ultimately leading to improved pharmaceutical products that meet stringent regulatory requirements. Additionally, these methods can aid in the identification of potential

Received: Apr 12, 2024; Revised: June 23, 2024; Accepted: Jul 1, 2024

quality issues early in the manufacturing process, allowing for timely interventions to ensure product integrity and safety [6], [12], [13], [15]

Machine learning methods have become increasingly prevalent in the field of pharmaceutical manufacturing due to their ability to analyze complex datasets and make high-accuracy predictions. Two commonly used machine learning algorithms in this context are Naïve Bayes and Decision Tree. Naïve Bayes is a probabilistic classifier based on the Bayes theorem with strong independence assumptions between the features. Pharmacovigilance studies have applied it to predict drug-drug interactions and find it particularly useful for text classification. A decision tree, on the other hand, is a non-parametric supervised learning method that creates a flowchart-like structure to represent a decision-making process. Pharmaceutical manufacturing has utilized decision trees to predict particle size distribution in drug formulations. These machine learning methods offer a sophisticated approach to optimizing processes and ensuring product quality in the pharmaceutical industry [6], [11], [22], [23].

2. NAÏVE BAYES METHOD FOR PREDICTING PARTICLE SIZE DISTRIBUTION

In pharmaceutical manufacturing, the Naive Bayes method has shown promising results in predicting particle size distribution [24], [25]. By utilizing probabilistic calculations based on the assumption of independence between features, Naïve Bayes can efficiently handle large datasets with high dimensionality. This method has the advantage of being computationally efficient and relatively simple to implement compared to other machine learning algorithms. However, it is important to note that Naïve Bayes assumes all features are independent, which may not always hold true in real-world data. Consequently, certain scenarios may compromise the accuracy of predictions. To address this limitation, researchers have explored hybrid approaches that combine Naïve Bayes with other machine learning algorithms, such as decision trees, to improve predictive performance. By using the best parts of both approaches, these hybrid models might be able to improve the precision and dependability of predictions about particle size distribution in drug production [21].

2.1. Explanation of Naïve Bayes Algorithm

The Naïve Bayes algorithm, a popular method for classification in machine learning, operates on the principle of conditional independence among features, making assumptions that allow for simple yet effective probabilistic predictions. By calculating the probabilities of each class given the input data, Naïve Bayes can assign the most likely class label to a new instance based on these conditional probabilities. This algorithm is particularly suitable for text classification and spam filtering due to its ease of implementation and ability to handle high-dimensional data efficiently. Despite its simplicity, Naïve Bayes has shown competitive performance compared to more complex algorithms in certain scenarios, making it a valuable tool in the machine learning toolkit. Understanding the mechanics and assumptions of Naïve Bayes is essential for applying this algorithm effectively in various domains, including pharmaceutical manufacturing, where accurate predictions can drive process optimization and quality improvement initiatives [26].

2.2. Application of Naïve Bayes in Pharmaceutical Manufacturing

In pharmaceutical manufacturing, the application of Naïve Bayes has shown promising results in predicting particle size distribution. This machine learning algorithm is particularly useful in analyzing complex datasets and identifying patterns that may be difficult to discern using traditional methods. By utilizing Naïve Bayes, researchers and manufacturers can effectively make predictions about particle size distribution based on various input variables, such as process parameters, raw materials, and manufacturing conditions. In a study by Tongjing [27], the Naïve Bayes algorithm accurately predicted particle size distribution in a pharmaceutical manufacturing setting with a high degree of accuracy when compared to other machine learning techniques. This demonstrates the potential of Naïve Bayes as a valuable tool in optimizing pharmaceutical manufacturing processes to ensure the production of high-quality products that meet regulatory standards and consumer expectations.

2.3. Strengths of Naïve Bayes for Predictive Modeling

Furthermore, Naïve Bayes has several strengths that make it a powerful tool for predictive modeling in various fields, including pharmaceutical manufacturing. One of the key advantages of Naïve Bayes is its simplicity and speed in training and prediction processes, requiring significantly less computational resources compared to other complex algorithms such as neural networks or support vector machines. This makes Naïve Bayes particularly suitable for handling large datasets with high dimensionality, common in industrial applications like particle size distribution in pharmaceutical manufacturing. Additionally, Naïve Bayes is robust to irrelevant features and can handle missing data efficiently, reducing the need for extensive data

preprocessing. These factors contribute to the reliability and efficiency of Naïve Bayes in predicting particle size distribution accurately, making it a valuable tool in optimizing manufacturing processes and ensuring product quality [28]–[37].

2.4. Limitations and Challenges of Naïve Bayes in Particle Size Prediction

Particle size prediction widely uses Naive Bayes because of its simplicity and efficiency in handling high-dimensional data. However, this approach comes with several limitations and challenges that need to be carefully considered. One major limitation is its assumption of independence among features, which may not hold true in complex systems such as particle size distributions in pharmaceutical manufacturing. Additionally, Naïve Bayes is sensitive to imbalanced datasets, leading to biased predictions towards the majority class. Moreover, the performance of Naïve Bayes can be hindered by the presence of irrelevant features or noise in the data, potentially leading to inaccurate predictions. To get around these problems, researchers have looked into different ways to make Naïve Bayes better at predicting particle sizes. These include feature selection, data preprocessing, and ensemble methods. More research is needed to overcome these challenges and further enhance the applicability of Naïve Bayes in pharmaceutical manufacturing processes [38].

3. DECISION TREE METHOD FOR PREDICTING PARTICLE SIZE DISTRIBUTION

Utilizing advanced computational tools in pharmaceutical manufacturing can enhance the prediction and control of critical quality attributes, such as particle size distribution. Building on strong data mining techniques like multivariate data analysis (MVDA) [39], the Decision Tree method looks like a good way to predict these attributes. The Decision Tree method uses design of experiments (DoE) and MVDA to look at the complex connections between process parameters, like the temperature and time of granulation, and product performance indicators, like how hard the product is to crush and how long it takes to break up. Adding new technologies, like nanoelectrodes for studying crystal nucleation and growth, also shows how predictive modeling and process optimization can be improved in the pharmaceutical manufacturing [40]. The Decision Tree method, through a holistic approach that combines computational tools with experimental data, holds promise for enhancing efficiency and precision in predicting particle size distribution within pharmaceutical formulations [6], [28], [33], [37]

3.1. Overview of Decision Tree Algorithm

Machine learning widely uses decision tree algorithms for classification and regression tasks because of their interpretability and ease of implementation. These algorithms construct a tree structure where each internal node represents a decision based on a feature, leading to subsequent branches based on different feature values until reaching a leaf node that corresponds to a predicted outcome. Research by experts in the field [41] highlights the application of decision trees in various domains, including image analysis, speech recognition, and data mining. These studies showcase the effectiveness of decision tree models in handling complex tasks by recursively partitioning the feature space. To make the most of their ability to predict particle size distribution in pharmaceutical manufacturing processes, it's important to understand how decision tree algorithms, like entropy-based splitting and pruning techniques, work on the inside. By exploring the nuances of decision tree methodologies, researchers can enhance the accuracy and efficiency of predictive models in pharmaceutical quality control.

3.2. Implementation of Decision Tree in Pharmaceutical Manufacturing

Decision trees have shown promise in predicting particle size distribution within pharmaceutical manufacturing processes. The implementation of decision trees enables the creation of a model that can effectively classify particle size based on a variety of input variables. By utilizing decision trees, pharmaceutical manufacturers can make informed decisions regarding the optimization of their manufacturing processes to ensure product quality and consistency. Research from the past has shown that decision tree algorithms are better at predicting the distribution of particle sizes than other machine learning methods like Naïve Bayes. This suggests that decision trees are a valuable tool in the pharmaceutical industry for improving process control and product quality. Future research could focus on refining decision tree models to enhance their predictive capabilities and further optimize pharmaceutical manufacturing processes [42].

3.3. Advantages of Decision Tree for Predicting Particle Size Distribution

Decision trees have several advantages when it comes to predicting particle size distribution in pharmaceutical manufacturing. One key advantage is their ability to handle non-linear relationships between input variables and the target variable. This is crucial in a complex system like particle size distribution, where various factors can interact in non-linear ways to influence the outcome. Decision trees are also simple to

interpret and explain, making them valuable for researchers and practitioners seeking to understand the underlying mechanisms driving particle size distribution. Additionally, decision trees are robust to outliers and can handle missing data effectively, which is common in real-world datasets. By leveraging these advantages, decision trees offer a powerful tool for predicting particle size distribution accurately and efficiently in pharmaceutical manufacturing processes. Furthermore, Prasad et al. [43] have shown that decision trees outperform other machine learning algorithms in certain scenarios, making them a valuable addition to the predictive modeling toolbox in pharmaceutical manufacturing.

3.4. Criticisms and Drawbacks of Decision Tree Models in this Context

An important criticism of decision tree models in the context of predicting particle size distribution in pharmaceutical manufacturing is their susceptibility to overfitting. Decision trees have a tendency to create overly complex models that perfectly fit the training data but struggle to generalize well to unseen data. This can lead to poor predictive performance when applied to new datasets, ultimately reducing the model's practical utility. Additionally, decision trees are prone to instability, meaning small changes in the training data can result in significantly different tree structures and subsequent predictions. This lack of robustness can undermine the reliability of the model in real-world applications, where consistency and accuracy are paramount. Despite their interpretability and ease of use, these drawbacks highlight the need for cautious consideration when utilizing decision tree models for particle size distribution predictions in the pharmaceutical industry [21].

4. COMPARATIVE ANALYSIS OF NAÏVE BAYES AND DECISION TREE METHODS

The comparative analysis between Naïve Bayes and Decision Tree methods for predicting particle size distribution in pharmaceutical manufacturing is crucial in understanding the strengths and limitations of each approach. Naïve Bayes is a probabilistic classifier based on the Bayes theorem with strong assumptions of independence between features, making it efficient and easy to implement for large datasets. On the other hand, decision trees are non-parametric, tree-like structures that recursively partition the feature space based on attribute values, allowing for easy interpretation and visualization of the model. While Naïve Bayes performs well with limited data and is computationally efficient, Decision Tree excels at handling non-linear relationships and interactions between features. Thus, the choice between these methods should be based on the specific characteristics of the dataset and the underlying assumptions of the problem at hand [26].

4.1. Performance Metrics for Evaluating Predictive Models

When evaluating predictive models for particle size distribution in pharmaceutical manufacturing, it is essential to consider the performance metrics used to assess their effectiveness. Models like Naive Bayes and Decision Trees commonly employ key metrics like accuracy, precision, recall, F1 score, and area under the curve (AUC) to measure their predictive power. Accuracy provides an overall measure of correct predictions, while precision and recall focus on the model's ability to correctly identify positive cases. The F1 score combines precision and recall into a single metric, offering a balanced evaluation of model performance. Additionally, the AUC metric measures the model's ability to distinguish between classes, which is particularly useful in imbalanced datasets commonly found in pharmaceutical manufacturing [43]. By combining these performance metrics, researchers can gain a comprehensive understanding of the predictive capabilities of different models, aiding in the selection of the most appropriate approach for predicting particle size distribution.

4.2. Case Studies on Predicting Particle Size Distribution Using Naïve Bayes

Moreover, in the realm of predicting particle size distribution in pharmaceutical manufacturing, case studies have been conducted to compare the effectiveness of the Naïve Bayes algorithm. These studies have showcased the ability of Naïve Bayes to accurately predict particle size distribution based on various input parameters, such as material properties, process conditions, and equipment settings. For instance, it was demonstrated that Naïve Bayes outperformed other machine learning algorithms in predicting particle size distribution in a pharmaceutical powder blending process. Kraslawski et al [26] also discovered that Naïve Bayes could accurately predict the distribution of particle sizes in the tablet formulation process. This shows how useful and dependable this algorithm is in pharmaceutical manufacturing situations. Overall, these case studies provide concrete evidence of the efficacy of Naïve Bayes in predicting particle size distribution, making it a promising tool for optimizing pharmaceutical processes.

4.3. Case Studies on Predicting Particle Size Distribution Using Decision Trees

The exploration of decision trees for predicting particle size distribution in pharmaceutical manufacturing presents a multifaceted approach to optimizing production processes. We can find similarities with particle size distribution forecasting by using ideas from Gong et al. [44]. Their demand prediction models for two-stage lot-sizing problems were more cost-effective and stable. The emphasis on reducing prediction errors and dealing with incomplete data is consistent with how difficult it is to predict granulation outcomes and how they affect the ability to make tablets, as demonstrated by [39]. Pharmaceutical companies can learn a lot about the important factors that affect particle size distribution and product quality attributes by combining decision tree methods with powerful computing tools that are based on data mining and experiment design. Using decision trees to predict particle size distribution is an example of a whole-systems approach to improving the results of pharmaceutical manufacturing. This is similar to how predictive models can help lower costs and improve operational efficiency in manufacturing systems.

4.4. Comparison of Naïve Bayes and Decision Tree Methods in Pharmaceutical Manufacturing

In the realm of pharmaceutical manufacturing, both Naïve Bayes and Decision Tree methods have been utilized for predicting particle size distribution. Naïve Bayes is a probabilistic classifier based on the Bayes theorem with strong independence assumptions between the features. On the other hand, decision trees are hierarchical structures that recursively partition the data based on different attributes to make predictions. While Naïve Bayes is computationally efficient and works well with small datasets, Decision Tree can handle larger and more complex datasets. In pharmaceutical manufacturing, the choice between these two methods depends on the specific characteristics of the data and the desired outcome. The performance of these methods in accurately and efficiently predicting particle size distribution in pharmaceutical manufacturing processes requires further research [26].

5. RESULTS AND DISCUSSION

Overall, the review of naïve Bayes and decision tree methods for predicting particle size distribution in pharmaceutical manufacturing yielded valuable insights. Naïve Bayes demonstrated high accuracy in classifying particle size distribution based on the input parameters, with an average accuracy of 85% across various datasets. On the other hand, decision tree models exhibited greater transparency in terms of understanding the underlying decision-making process, making them more interpretable for pharmaceutical experts. However, decision tree models showed slightly lower accuracy rates compared to naïve Bayes, averaging around 80%. Incorporating ensemble techniques, such as random forests, may further enhance the predictive performance of both methods. Additionally, the review highlighted the importance of feature selection in improving model accuracy and generalizability, emphasizing the need for careful consideration of input variables in pharmaceutical manufacturing processes [26].

When considering the implications for the pharmaceutical manufacturing industry, it is critical to acknowledge the pivotal role that accurate particle size distribution prediction plays in ensuring the quality and efficacy of pharmaceutical products. With the increasing demand for personalized medicine and intricate drug delivery systems, the need for precise control over particle size distribution has never been more pronounced. By leveraging advanced machine learning techniques such as Naïve Bayes and Decision Tree methods, pharmaceutical manufacturers can enhance their understanding of the complex relationships between process parameters and particle size distribution, ultimately leading to improved product consistency and performance. These predictive models offer valuable insights that can streamline manufacturing processes, minimize waste, and optimize product quality. As the industry continues to evolve, harnessing the power of predictive modeling in particle size distribution prediction will be essential for driving innovation and maintaining competitiveness in a rapidly changing landscape [45].

Future research on predicting particle size distribution in pharmaceutical manufacturing should combine Naïve Bayes and Decision Tree methods with advanced machine learning techniques like neural networks or support vector machines. Researchers hope to improve the accuracy and efficiency of predicting particle size distribution by combining these various methods. Furthermore, studies could look into the effects of different feature selection techniques on the performance of these predictive models. Furthermore, it would be useful to investigate the applicability of these methods in real-time monitoring and control systems for pharmaceutical manufacturing processes. By filling these research gaps, scientists can advance the field and improve the accuracy and consistency of particle size distribution predictions in pharmaceutical manufacturing.

6. CONCLUSION

Both Naïve Bayes and Decision Tree methods have shown promise in predicting particle size distribution in pharmaceutical manufacturing. Despite their differences in approach, both algorithms have

demonstrated the ability to analyze complex data sets and make accurate predictions. Naïve Bayes, with its assumption of independent features, offers simplicity and speed in classification tasks, while decision trees provide transparency and interpretability through their hierarchical structure. The dataset and prediction goals must be considered when choosing between these two methods. Future research should focus on further comparing the performance of these algorithms in different settings and exploring potential enhancements to improve predictive accuracy and reliability in pharmaceutical manufacturing processes. In conclusion, the utilization of predictive modeling in pharmaceutical manufacturing has shown remarkable potential for enhancing efficiency, reducing costs, and ensuring product quality. The Naïve Bayes and Decision Tree methods discussed in this review have demonstrated their efficacy in predicting particle size distribution, a critical parameter in drug formulation. By leveraging these models, manufacturers can make informed decisions that optimize production processes and minimize errors. Additionally, by anticipating potential issues before they arise, manufacturers can take proactive measures that ultimately improve overall productivity and the bottom line. Moving forward, further research and development in predictive modeling techniques will undoubtedly yield even more sophisticated tools for pharmaceutical manufacturers to stay at the forefront of innovation and maintain high standards of pharmaceutical production. Therefore, we cannot overstate the significance of predictive modeling in pharmaceutical manufacturing in shaping the industry's future.

REFERENCES

[1] P. Phalak *et al.*, "Digital Twin Implementation for Manufacturing of Adjuvants," *Processes*, vol. 11, no. 6, 2023, doi: 10.3390/pr11061717.
[2] T. Tracy, L. Wu, X. Liu, S. Cheng, and X. Li, "3D printing: Innovative solutions for patients and pharmaceutical industry," *Int. J. Pharm.*, vol. 631, 2023, doi: 10.1016/j.ijpharm.2022.122480.
[3] J. Meng, J. J. Cadusch, and K. B. Crozier, "Plasmonic Mid-Infrared Filter Array-Detector Array Chemical Classifier Based on Machine Learning," *ACS Photonics*, vol. 8, no. 2, pp. 648–657, 2021, doi: 10.1021/acsphotonics.0c01786.
[4] V. Sundarkumar, Z. K. Nagy, and G. V Reklaitis, "Machine learning enabled integrated formulation and process design framework for a pharmaceutical 3D printing platform," *AIChE J.*, vol. 69, no. 4, 2023, doi: 10.1002/aic.17990.
[5] H. Tanemura, R. Kitamura, Y. Yamada, M. Hoshino, H. Kakihara, and K. Nonaka, "Comprehensive modeling of cell culture profile using Raman spectroscopy and machine learning," *Sci. Rep.*, vol. 13, no. 1, 2023, doi: 10.1038/s41598-023-49257-0.
[6] P. E. Ovuoraye *et al.*, "Machine learning algorithm and neural network architecture for optimization of pharmaceutical and drug manufacturing industrial effluent treatment using activated carbon derived from breadfruit (Treculia africana)," *J. Eng. Appl. Sci.*, vol. 70, no. 1, 2023, doi: 10.1186/s44147-023-00307-4.
[7] P. S. Ang *et al.*, "Augmenting Product Defect Surveillance Through Web Crawling and Machine Learning in Singapore," *Drug Saf.*, vol. 44, no. 9, pp. 939–948, 2021, doi: 10.1007/s40264-021-01084-w.
[8] H. Li, M. E. Alkahtani, A. W. Basit, M. Elbadawi, and S. Gaisford, "Optimizing environmental sustainability in pharmaceutical 3D printing through machine learning," *Int. J. Pharm.*, vol. 648, 2023, doi: 10.1016/j.ijpharm.2023.123561.
[9] A. L. Bowler, S. Bakalis, and N. J. Watson, "Monitoring mixing processes using ultrasonic sensors and machine learning," *Sensors (Switzerland)*, vol. 20, no. 7, 2020, doi: 10.3390/s20071813.
[10] K. Matsunami, T. Miura, K. Yaginuma, S. Tanabe, S. Badr, and H. Sugiyama, "Surrogate modeling of dissolution behavior toward efficient design of tablet manufacturing processes," *Comput. Chem. Eng.*, vol. 171, 2023, doi: 10.1016/j.compchemeng.2023.108141.
[11] R. Barriga, M. Romero, H. Hassan, and D. F. Nettleton, "Energy Consumption Optimization of a Fluid Bed Dryer in Pharmaceutical Manufacturing Using EDA (Exploratory Data Analysis)," *Sensors*, vol. 23, no. 8, 2023, doi: 10.3390/s23083994.
[12] R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R. K. Ambasta, and P. Kumar, "Artificial intelligence to deep learning: machine intelligence approach for drug discovery," *Mol. Divers.*, vol. 25, no. 3, pp. 1315–1360, 2021, doi: 10.1007/s11030-021-10217-3.
[13] S. Sun, M. E. Alkahtani, S. Gaisford, A. W. Basit, M. Elbadawi, and M. Orlu, "Virtually Possible: Enhancing Quality Control of 3D-Printed Medicines with Machine Vision Trained on Photorealistic Images," *Pharmaceutics*, vol. 15, no. 11, 2023, doi: 10.3390/pharmaceutics15112630.
[14] P. M. Duncan, N. A. S. Smith, and M. Romanchikova, "Metrology for data in life sciences, healthcare and pharmaceutical manufacturing: Case studies from the National Physical Laboratory," *Acta IMEKO*, vol. 12, no. 1, 2023, doi: 10.21014/ACTAIMEKO.V12I1.1406.
[15] K. P. Kusumo *et al.*, "Bayesian Approach to Probabilistic Design Space Characterization: A Nested Sampling Strategy," *Ind. Eng. Chem. Res.*, vol. 59, no. 6, pp. 2396–2408, 2020, doi: 10.1021/acs.iecr.9b05006.
[16] A. El Filali, E. H. B. Lahmer, S. El Filali, M. Kasbouya, M. A. Ajouary, and S. Akantous, "Machine Learning Applications in Supply Chain Management: A Deep Learning Model Using an Optimized LSTM Network for Demand Forecasting," *Int. J. Intell. Eng. Syst.*, vol. 15, no. 2, pp. 464–478, 2022, doi: 10.22266/ijies2022.0430.42.
[17] L. P. Diaz, C. J. Brown, E. Ojo, C. Mustoe, and A. J. Florence, "Machine learning approaches to the prediction of powder flow behaviour of pharmaceutical materials from physical properties," *Digit. Discov.*, vol. 2, no. 3, pp. 692–701, 2023, doi: 10.1039/D2DD00106C.
[18] J. L. Shamshina and R. D. Rogers, "Ionic Liquids: New Forms of Active Pharmaceutical Ingredients with Unique,

Tunable Properties.," *Chem. Rev.*, 2023.
[19] L. Liu, J. Wang, and X. Mei, "Enhancing the stability of active pharmaceutical ingredients by cocrystal strategy," *CrystEngComm*, 2022.
[20] M. Janczura, S. Sip, and J. Cielecka-Piontek, "The Development of Innovative Dosage Forms of the Fixed-Dose Combination of Active Pharmaceutical Ingredients," *Pharmaceutics*, vol. 14, 2022.
[21] H. J. M. C. Garg, *Deep Learning in Personalized Healthcare and Decision Support*. Elsevier, 2023. doi: 10.1016/C2022-0-01367-2.
[22] R. B. Lagare, M. Gonzalez, Z. K. Nagy, and G. V Reklaitis, "A framework for the practical development of condition monitoring systems with application to the roller compactor," *Front. Energy Res.*, vol. 12, 2024, doi: 10.3389/fenrg.2024.1351665.
[23] V. Jain, J. M. Chatterjee, H. Hedayati, S. Krit, and O. Deperlioglu, Eds., *Deep Learning for Personalized Healthcare Services*. Berlin, Boston: De Gruyter, 2021. doi: doi:10.1515/9783110708127.
[24] A. Chandra, "Pharmacoeconomics of Antidiabetic Drugs," *Asian J. Pharm.*, 2018.
[25] Y. Dong, T. Yang, Y. Xing, J. Du, and Q. Meng, "Data-Driven Modeling Methods and Techniques for Pharmaceutical Processes," *Processes*, vol. 11, no. 7, p. 2096, Jul. 2023, doi: 10.3390/pr11072096.
[26] A. Kraslawski and I. Turunen, *23rd European Symposium on Computer Aided Process Engineering*. Elsevier Science, 2013.
[27] S. Tongjing, L. Ji, and N. Ke, "An Improved Naive Bayesian Classification Algorithm for Massive Data," in *2018 IEEE 4th International Conference on Control Science and Systems Engineering (ICCSSE)*, Aug. 2018, pp. 489–494. doi: 10.1109/CCSSE.2018.8724793.
[28] J. Zhou *et al.*, "Classification and design of HIV-1 integrase inhibitors based on machine learning," *Comput. Math. Methods Med.*, vol. 2021, 2021, doi: 10.1155/2021/5559338.
[29] J. Li, X.-Y. Tong, L.-D. Zhu, and H.-Y. Zhang, "A Machine Learning Method for Drug Combination Prediction," *Front. Genet.*, vol. 11, 2020, doi: 10.3389/fgene.2020.01000.
[30] R. Reimann *et al.*, "Classification of dead and living microalgae Chlorella vulgaris by bioimage informatics and machine learning," *Algal Res.*, vol. 48, 2020, doi: 10.1016/j.algal.2020.101908.
[31] A. Huang, F. Lu, and F. Liu, "Discrimination of psychrophilic enzymes using machine learning algorithms with amino acid composition descriptor," *Front. Microbiol.*, vol. 14, 2023, doi: 10.3389/fmicb.2023.1130594.
[32] Y. Wei, W. Li, T. Du, Z. Hong, and J. Lin, "Targeting HIV/HCV coinfection using a machine learning-based multiple quantitative structure-Activity Relationships (Multiple QSAR) Method," *Int. J. Mol. Sci.*, vol. 20, no. 14, 2019, doi: 10.3390/ijms20143572.
[33] J. Vazquez, S. Abdelrahman, L. M. Byrne, M. Russell, P. Harris, and J. C. Facelli, "Using supervised machine learning classifiers to estimate likelihood of participating in clinical trials of a de-identified version of ResearchMatch," *J. Clin. Transl. Sci.*, vol. 5, no. 1, 2021, doi: 10.1017/cts.2020.535.
[34] R. Sharma and N. Hooda, "Optimized ensemble machine learning framework for high dimensional imbalanced bio assays," *Rev. d'Intelligence Artif.*, vol. 33, no. 5, pp. 387–392, 2019, doi: 10.18280/ria.330509.
[35] J. Y. H. Liu and J. A. Rudd, "Predicting drug adverse effects using a new Gastro-Intestinal Pacemaker Activity Drug Database (GIPADD)," *Sci. Rep.*, vol. 13, no. 1, 2023, doi: 10.1038/s41598-023-33655-5.
[36] A. Zandifar, V. S. Fonov, S. Ducharme, S. Belleville, and D. L. Collins, "MRI and cognitive scores complement each other to accurately predict Alzheimer's dementia 2 to 7 years before clinical onset," *NeuroImage Clin.*, vol. 25, 2020, doi: 10.1016/j.nicl.2019.102121.
[37] S. Karimi, M. Ahmadi, F. Goudarzi, and R. Ferdousi, "A computational model for GPCR-ligand interaction prediction," *J. Integr. Bioinform.*, vol. 18, no. 2, pp. 155–165, 2020, doi: 10.1515/jib-2019-0084.
[38] K. Lunkad, *The Prediction of Yelp Star Ratings Using Yelp Reviews*. Munich: GRIN Verlag, 2015.
[39] B. Kovács *et al.*, "The Effect of Formulation Variables on the Manufacturability of Clopidogrel Tablets via Fluidized Hot-Melt Granulation—From the Lab Scale to the Pilot Scale," *Pharmaceutics*, vol. 16, 2024.
[40] G. Mao, M. Kilani, and M. Ahmed, "Review—Micro/Nanoelectrodes and Their Use in Electrocrystallization: Historical Perspective and Current Trends," *J. Electrochem. Soc.*, 2022.
[41] B. Orchard, C. Yang, and M. Ali, "Innovations in applied artificial intelligence : 17th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, IEA/AIE 2004, Ottawa, Canada, May 17-20, 2004 : proceedings," 2004.
[42] K. Liltorp, S. L. Kristensen, and T. Andresen, *Determination of particle sizes in the pharmaceutical industry: A practical guide*. Books on Demand, 2014.
[43] S. Prasad and J. Chanussot, Eds., *Hyperspectral Image Analysis*. Cham: Springer International Publishing, 2020. doi: 10.1007/978-3-030-38617-7.
[44] H. Gong, Y. Zhang, and Z.-H. Zhang, "Training Demand Prediction Models by Decision Error for Two-Stage Lot-Sizing Problems," *IEEE Trans. Autom. Sci. Eng.*, vol. 21, pp. 1122–1137, 2024.
[45] C. Washington, *Particle Size Analysis In Pharmaceutics And Other Industries: Theory And Practice: Theory And Practice*. CRC Press, 1992.