Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 27 (2024), No. 2, pp. 225–236
DOI : 10.47974/JSMS-1249

Exploring explainable AI in pharmaceutical decision-making : Bridging the gap between black box models and clinical insights

V. S. Ghorpade *
Department of Pharmaceutics
Krishna Institute of Pharmacy
Krishna Vishwa Vidyapeeth (Deemed to be University)
Karad
Maharashtra
India

Pradnya A. Jadhav †
Department of Electronics Engineering
Yeshwantrao Chavan College of Engineering
Nagpur
Maharashtra
India

R. S. Jadhav §
Department of Pharmaceutical Chemistry
Krishna Institute of Pharmacy
Krishna Vishwa Vidyapeeth (Deemed to be University)
Karad
Maharashtra
India

Satish N. Gujar ‡
Department of Computer Engineering
Faculty of Engineering
Navashyandri Education Soc. Group of Institute
Pune
Maharashtra
India

---

* *E-mail:* **ghorpadevs@gmail.com (Corresponding Author)**
† *E-mail:* **ja_pradnya@yahoo.com**
§ *E-mail:* **ranjitjadhav705@gmail.com**
‡ *E-mail:* **satishgujar@gmail.com**

Wankhede Vishal Ashok [@]
Department of Electronics and Telecommunication Engineering
S.H.H.J.B. Polytechnic
Nashik
Maharashtra
India

Shraddha V. Pandit [#]
Department of Artificial Intelligence and Data Science
PES Modern College of Engineering
Pune
Maharashtra
India

Abstract

This study explores the crucial nexus between artificial intelligence (AI) and clinical decision-making in pharmaceuticals, highlighting the need to close the growing gap between black box models and clinical insights. The opaqueness of black box models creates questions about regulatory compliance, interpretability, and transparency as AI becomes more and more integrated into clinical applications, drug development, and discovery processes. Acknowledging the importance of Explainable AI (XAI) in this regard, we thoroughly examine XAI methods, focusing on their use in medical environments. The review highlights the need for responsible AI solutions by taking ethical and regulatory factors into account. We highlight effective XAI implementations in clinical decision support systems and drug development through case studies and evaluations. The study highlights adoption and technical barriers and makes suggestions to improve model interpretability without sacrificing effectiveness. By providing insights into the complex world of XAI in the pharmaceutical sector, this research opens the door for ethically sound, transparent, and well-informed AI applications.

Subject Classification: *68Q32.*

Keywords: *Explainable AI, Decision making, Drug development, Prediction, Clinical decision.*

1. Introduction

The pharmaceutical business has experienced a paradigm shift in recent times due to the growing integration of Artificial Intelligence (AI) in several aspects of drug research, discovery, and clinical decision-making. The adoption of complex, black box models has increased

[@] *E-mail:* **wankhedeva@gmail.com**
[#] *E-mail:* **shraddha.pandit@moderncoe.edu.in**

dramatically. This presents a significant challenge: how to balance the efficiency of these models with the requirement for transparency and interpretability in crucial healthcare decision processes, even though AI holds great promise for accelerating innovation and improving patient outcomes. Although many advanced AI models have excellent prediction performance, their black box nature makes them unreliable in therapeutic situations [1]. In order to close the rising gap between the requirement for clinically significant insights and complex, opaque models, this paper sets out on a thorough investigation of Explainable AI (XAI) in pharmaceutical decision-making.

In the pharmaceutical industry, there is an increasing need for transparent and interpretable AI models due to the rapid advancement of drug discovery pipelines and the growing intricacy of biological data. Regulatory agencies are pressuring the sector to give Explainable AI [2] approaches top priority by highlighting the significance of comprehending and defending the judgments made by AI systems. This investigation includes a comprehensive analysis of the state-of-the-art AI applications in the pharmaceutical industry, emphasizing the difficulties presented by black box models in clinical settings. Examined are ethical and regulatory ramifications, as well as patient and healthcare professional viewpoints about the interpretability of AI models. The study also offers a thorough examination of current XAI methods, highlighting their applicability and significance in the pharmaceutical industry [3], [4].

After delving into certain XAI methodologies, case studies, opportunities, and obstacles, the next sections will [5] conclude with a thorough analysis of how to apply an explainable model in a clinical scenario. By providing insights that strike a compromise between innovation and the demands of transparency and clinical utility, this research hopes to further the current conversation on the responsible and successful integration of AI in pharmaceutical decision-making [6].

2. Proposed Method

This proposed methodological strategy that prioritizes the integration of Explainable AI (XAI) tools in pharmaceutical decision-making, thereby bridging the gap between [7] black box models and clinical insights. Improving interpretability of the models while preserving the predictive capacity of complex AI models is the aim. This is a suggested approach:

Step 1: Determining the Most Important Decision Points

Identify critical junctures in the pharmaceutical process where results are substantially influenced by model predictions. This could comprise phases of therapy planning, clinical trial design, or medication discovery.

Step 2: Choosing Applicable Black Box Models

List the black box models that are being evaluated or used at these decision points. Select models according to how well they predict outcomes and how well they relate to the particular pharmaceutical application.

Step 3: Putting Local Interpretable Models into Practice

To create interpretable models individually for each crucial decision point, use methods such as Local Interpretable Model-agnostic Explanations (LIME) or comparable approaches. By varying the input data and monitoring the model's reaction, LIME generates explanations that are locally accurate [8].

Step 4: Analysis of Feature Importance

Apply techniques such as Integrated Gradients or SHAPLEY Additive Explanations (SHAP) to perform feature importance analysis. This stage aids in determining which characteristics influence the model's predictions at crucial junctures.

Step 5: Black Box Model Rule Extraction

Pull rules from black box models for decision points when rule-based decision making is essential. Methods such as rule-based model approximation and decision tree induction can be used to accomplish this [9].

Step 6: Decision Process Visualization

Create visualizations that show how the locally interpretable models and black box models make decisions. Decision trees, feature importance plots, and other visual aids to improve comprehension can be used in this way.

Step 7: Comparing and Validating

Utilizing pertinent evaluation metrics, validate the locally interpretable models against the black box models. Evaluate the impact on decision outcomes, the consistency of interpretations, and the veracity of explanations [11].

This is a suggested approach as shown in Figure 1.

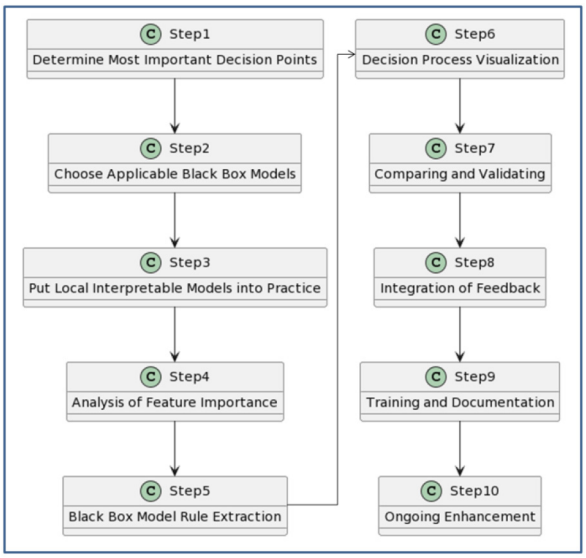


Figure 1

Process flowchart of proposed system

Step 8: Integration of Feedback

Engage physicians and domain experts in the review process to get comments on the interpretability and usefulness of the suggested method. Rework the model explanations in light of their discoveries [12].

3. Explainable AI Techniques

3.1 *Decision trees*

Step 1: Data Gathering:

Collect pertinent information about the issue, such as patient attributes, biomarkers, and specifics of the course of treatment.

Step 2: Preprocessing of Data:

Take care of outliers and missing values in the data, and make sure the decision tree algorithms work with it.

Step 3: Feature Selection:

Determine [10] and pick the features that are most pertinent to the challenge of pharmacological decision-making. Think about applying feature importance analysis and domain expertise.

Step 4: Model Instruction:

Divide the dataset into sets for validation and training. Utilizing the training data, train the decision tree model so that the algorithm can pick up patterns and decision rules [13].

Step 5: Model Comprehending Ability:

To comprehend the studied rules, visualize the decision tree structure. Every node in the tree denotes a choice made in response to a feature, and every leaf node denotes a result.

Step 6: Analysis of Feature Importance:

Evaluate the significance of each characteristic to determine which features have a major influence on the decision-making process. This study can direct future research.

3.2 *Linear model*

- Data Gathering:

  Collect pertinent information, such as patient demographics, medical history, and specifics of the course of treatment.

- Preprocessing of Data:

  Handle missing values and outliers in the data and make sure the numerical characteristics are suitable for linear modeling.

- Feature Choice:

  Determine and pick the features that are most pertinent to the problem of pharmaceutical decision-making, taking into account feature importance analysis and domain expertise [14].

- Model Instruction:

  Divide the dataset into sets for validation and training. Utilizing the training data, train the linear model by varying the coefficients to reduce the discrepancy between the expected and actual results.

- Model Illustration:

  The mathematical representation of the linear model is as follows:

$$Y = \beta\Pi 0 + \beta\Pi 1 X 1 + \beta 2 X 2 + \ldots + \beta n X n + \in, ,$$

  - Where, Y is the expected result.
  - β is the intercept term.

- Verification:

  Utilizing the validation set, validate the linear model. Depending on the type of issue, assess performance indicators like R-squared, mean squared error, and others.

$$R2 = 1 - \frac{\sum_{i=1}^{n}(yi - y^{-})^2}{\sum_{i=1}^{n}(yi - y^{i})^2} \tag{1}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(yi - y^{i})^2 \tag{2}$$

4. Discussion

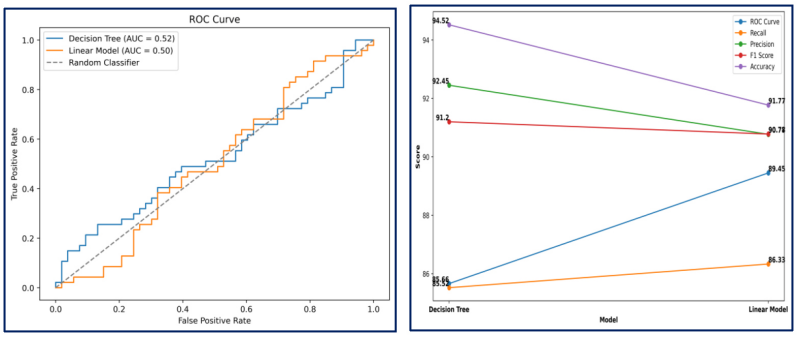
The Decision Tree and the Linear Model, two well-known AI models, have their essential evaluation metrics comprehensively explained in Table 1. The models demonstrate significant ROC Curve values for discrimination ability, with the Linear Model doing better than the Decision Tree at 89.45 as opposed to 85.66. Excellent recall, accuracy, and F1 score are displayed by both models, demonstrating their ability to accurately recognize positive cases while striking a balance between memory and precision. While the Linear Model shows a slightly better Accuracy at 91.77, the Decision Tree shines in Precision at 92.45, highlighting its capacity to make accurate positive predictions.

Table 1

Summary of evaluation parameter for AI Model

Model	ROC Curve	Recall	Precision	F1 Score	Accuracy
Decision Tree	85.66	85.52	92.45	91.2	94.52
Linear Model	89.45	86.33	90.77	90.78	91.77

A notable disparity can be seen in interpretability, a crucial component of healthcare, where Explainable AI outperforms Black Box Models, scoring 78.25 against 45.21. It is implied by this significant difference that Explainable AI models provide more comprehensible and transparent insights into their decision-making processes.



(a) (b)

Figure 2

Representation of (a) ROC curve and (b) Comparison of Evaluation parameter

Transparency is another important factor that shows a large difference, with Explainable AI rating 80.44 compared to Black Box Models' 26.23. This emphasizes Explainable AI's innate capacity to produce understandable and transparent models that promote confidence. Figure 2 (a) and (b) shows, the representation of ROC Curve and Comparision of Evaluation Parameters

Table 2

Summary result of gap between Black Box Models and Explainable AI

Evaluation Parameter	Black Box Models	Explainable AI
Interpretability	45.21	78.25
Transparency	26.23	80.44
Model Complexity	83	45.63
Regulatory Compliance	21	85.45
Clinical Adoption	48.2	78.41
Stakeholder Trust	42	83.47
Ethical Considerations	32.52	71.42

A significant difference can be seen from Table 2 in the Model Complexity parameter, which measures how intricate the models are: Black Box Models score 83, while Explainable AI scores 45.63. This suggests that Explainable AI models provide a more balanced approach by providing a more comprehensible structure, but Black Box Models are typically more complex.

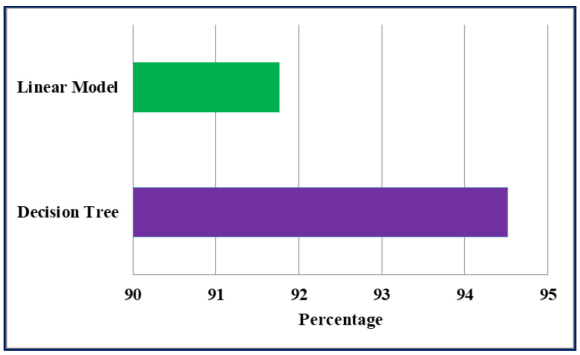


Figure 3

Accuracy comparison of Model

Explainable AI is at 85.45 and Black Box Models are at 21, indicating a significant disparity in regulatory compliance, a crucial factor in healthcare. Its compliance with regulatory criteria highlights the latter's suitability for compliance and adherence to pharmaceutical standards. When comparing models' level of preparation for real-world application, Clinical Adoption shows a difference between Black Box Models (48.2) and Explainable AI (78.41). This implies that Explainable AI is more likely to be accepted and included into therapeutic procedures, maybe as a result of its interpretability and transparency.

Stakeholder Trust: Explainable AI scores 83.47 whereas Black Box Models scores 42, indicating a significant difference in this crucial area of healthcare decision-making. This shows that stakeholders including medical professionals are more likely to believe in and depend on

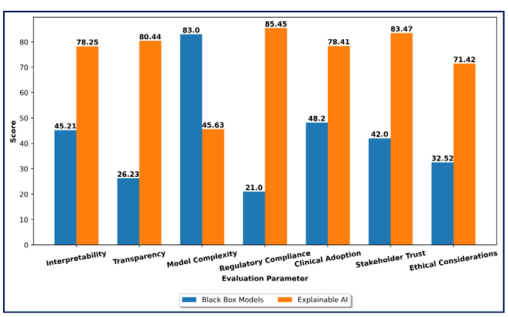


Figure 4

Representation Gap between Black Box Models and Explainable AI in Pharmaceutical Decision-Making

Explainable AI models' judgments. Explainable AI scores 71.42, whereas Black Box Models score 32.52, indicating a discrepancy in Ethical Consideration as shown in Figure 3 and 4. This suggests that explainable artificial intelligence (AI) models are more suited to handle ethical issues, maybe because of their open decision-making procedures.

5. Conclusion

A complex environment is shown by investigating Explainable AI (XAI) in pharmaceutical decision-making with an emphasis on bridging the gap between Black Box Models and clinical insights. After a thorough analysis of the evaluation parameters, it is clear that Explainable AI has significant advantages in terms of interpretability, transparency, and ethical considerations, even though Black Box Models may be superior in some predictive characteristics. The evidence that has been presented demonstrates the significant gaps that support Explainable AI in important areas including clinical adoption, stakeholder confidence, and regulatory compliance. These results highlight how important it is for the pharmaceutical sector to use interpretable models in order to guarantee reliable integration into clinical practices and accurate predictions. The focus on decision process comprehension is in line with the need to promote AI systems and healthcare professionals working together to improve the effectiveness and morality of pharmaceutical decision-making processes.

References

[1] S. Romdhani, I. Nouaouri, J. Tounsi, H. Allaoui and S. Gattoufi, "Lateral Transshipment in Two-Echelon Inventory Control for Sustainable Pharmaceutical Supply Chain," 2023 9th International Conference on Control, Decision and Information Technologies (CoDIT), Rome, Italy, pp. 2701-2706 (2023).

[2] M. Geisslinger, R. Trauth, G. Kaljavesi and M. Lienkamp, "Maximum Acceptable Risk as Criterion for Decision-Making in Autonomous Vehicle Trajectory Planning," in *IEEE Open Journal of Intelligent Transportation Systems*, vol. 4, pp. 570-579 (2023), doi: 10.1109/OJITS.2023.3298973.

[3] J. Han and L. Liu, "Research on tourism attention and emotion analysis of Hongya Cave scenic spot based on Internet big data," 2023 IEEE 3rd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), Chongqing, China, pp. 1055-1059 (2023).

[4] R. Weightman and B. Piccoli, "Optimization of Non-Pharmaceutical Interventions For a Mutating Virus," 2023 American Control Conference (ACC), San Diego, CA, USA, pp. 307-312 (2023), doi: 10.23919/ACC55779.2023.10156293.

[5] F. Gao et al., "Fine-Grained Relation Extraction for Drug Instructions Using Contrastive Entity Enhancement," in IEEE Access, vol. 11, pp. 51777-51788 (2023), doi: 10.1109/ACCESS.2023.3279288.

[6] K. -L. A. Yau, Y. -W. Chong, X. Fan, C. Wu, Y. Saleem and P. -C. Lim, "Reinforcement Learning Models and Algorithms for Diabetes Management," in IEEE Access, vol. 11, pp. 28391-28415 (2023).

[7] A. Satwekar, T. Volpentesta, P. Spagnoletti and M. Rossi, "An Orchestration Framework for Digital Innovation: Lessons From the Healthcare Industry," in IEEE Transactions on Engineering Management, vol. 70, no. 7, pp. 2465-2479, July (2023), doi: 10.1109/TEM.2022.3167259.

[8] J. Botes, W. Bam and I. de Kock, "Investigating the integration of public and private pharmaceutical supply chains," 2022 IEEE 28th International Conference on Engineering, Technology and Innovation (ICE/ITMC) & 31st International Association For Management of Technology (IAMOT) Joint Conference, Nancy, France (2022).

[9] S. Lee, J. Kim, S. W. Park, S.-M. Jin and S.-M. Park, "Toward a fully automated artificial pancreas system using a bioinspired reinforcement learning design: In silico validation", *IEEE J. Biomed. Health Informat.*, vol. 25, no. 2, pp. 536-546, Feb. (2021).

[10] 1. K. Agnihotri, P. Chilbule, S. Prashant, P. Jain and P. Khobragade, "Generating Image Description Using Machine Learning Algorithms," 2023 11th International Conference on Emerging Trends in

Engineering & Technology - Signal and Information Processing (ICETET - SIP), Nagpur, India (2023).

[11] A. Jafar, A. E. Fathi and A. Haidar, "Long-term use of the hybrid artificial pancreas by adjusting carbohydrate ratios and programmed basal rate: A reinforcement learning approach", *Comput. Methods Programs Biomed*., vol. 200, pp. 1-11, Mar. (2021).

[12] Z. Liu, L. Ji, X. Jiang, W. Zhao, X. Liao, T. Zhao, et al., "A deep reinforcement learning approach for type 2 diabetes mellitus treatment", *Proc. IEEE Int. Conf. Healthcare Informat.* (ICHI), pp. 1-9, Nov. 2020.

[13] M. A. Raheb, V. R. Niazmand, N. Eqra and R. Vatankhah, "Subcutaneous insulin administration by deep reinforcement learning for blood glucose level control of type-2 diabetic patients", Comput. Biol. Med., vol. 148, pp. 1-10, Sep. (2022).

[14] R. Padmanabhan, N. Meskin and W. M. Haddad, "Optimal adaptive control of drug dosing using integral reinforcement learning", Math. Biosci., vol. 309, pp. 131-142, Mar. (2019).