RESEARCH PAPER

A Scalable Machine Learning Framework for Predictive Pharmacokinetics and Drug Response Modeling

K. Kalaiselvi*[1], B. Satyanarayana Murthy[2], K.V.Y. Radha Sri[3], Ch. L. N. Deepika[4]

[1] Department of AI & ML, SRM Institute of Science and Technology, Ramapuram, Chennai – 600 089, Tamil Nadu, India

Email: kalaisek5@srmist.edu.in

[2] Department of Computer Science and Engineering, Bonam Venkata Chalamayya Engineering College, Odalarevu, Andhra Pradesh, India

Email: murthy2007.b@gmail.com

[3] Department of Computer Science and Engineering, Aditya University, Surampalem, Andhra Pradesh, India

Email: venkatayaminiradhasrik@adityauniversity.in

[4] Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

Email: ldeepu2474@gmail.com

Corresponding Author: kalaisek5@srmist.edu.in

Abstract

Predictive pharmacokinetics and drug response modeling are indispensable elements of precision medicine that make it possible to design personalized treatment protocols based on the analysis of interactions between patients and drugs using computational approaches. In this paper, we present a scalable machine learning algorithm for predictive pharmacokinetics and drug response modeling from the Genomics of Drug Sensitivity in Cancer (GDSC) dataset. Our framework employs multi-omics data, including genetic information, drug-related attributes, and other factors involved in the process of interaction, to simulate complex relationships between drugs and biological systems. To increase the robustness and improve the performance of our learning model in the presence of noise and in multidimensional biomedical spaces, we apply the Noise-Resilient Quantum Neural Network (NR-QNN). Our scalable algorithm is preprocessed, feature-extracted, and dimensionally-reduced to achieve better learning capabilities. Experiments showed that our framework outperforms state-of-the-art ML, CNN, and LSTM methods in predictive pharmacokinetics and drug response modeling with an accuracy of 98.2%, precision of 97.5%, F1 score of 96.8%, and RMSE of 1.035 while reducing computation time. We conclude that our NR-QNN-based scalable learning algorithm performs well in modeling nonlinear pharmacokinetic properties and predicting drug responses.

Keywords: Drug, biological, Noise-Resilient Quantum Neural Network, feature-extracted

How to cite this article: Kalaiselvi K, Murthy BS, Radha Sri KVY, Deepika CLN. A Scalable Machine Learning Framework for Predictive Pharmacokinetics and Drug Response Modeling. Int J Drug Deliv Technol. 2026;16(31s):274-278. DOI: 10.25258/ijddt.16.31s.34

1. Introduction

Predictions in pharmacokinetics and drug response modelling appear to be among those elements that make up the foundation of the new field of precision medicine [1]. In traditional methods of pharmacokinetic studies, there is quite frequent reliance on averaging values obtained from multiple patients as opposed to the parameters peculiar to a particular person [2]. At the same time, as the amount of clinical data accumulated in electronic medical records and omics databases increases quickly, it becomes necessary to construct computational models capable of adequately taking into account intricate interactions between drugs and the body. One of such advanced technologies is represented by ML algorithms [4].

In such scenario, scalable machine learning models can be considered a viable option towards building up integrative models based on heterogeneous pharmacology data and making personalized predictions of drug responses [5]. With the application of sophisticated algorithms like ensemble learning, neural nets, and time-series models, ML models can incorporate complex variations in the behavior of drugs at different times and among patients. With the help of feature extraction, dealing with high-dimensional data and optimization techniques, pharmacokinetic models

using machine learning approaches can improve prediction and clinical decision-making.

Recent Works

James et al.[6] have developed a range of machine learning and AI-based frameworks for predictive pharmacokinetics and drug response modeling to enhance personalized medicine and dosage optimization. These studies commonly use clinical datasets, electronic health records, and pharmacological databases, applying techniques such as deep learning, ensemble methods, and regression models for prediction tasks. However, most of the existing approaches face limitations such as low scalability, reduced generalization across diverse patient groups, and inadequate handling of complex and heterogeneous biomedical data.

Golriz et al.[7] Have developed ML-based frameworks that utilize pathway signatures to assess dysregulation of biological processes for predicting drug response in individual patients. These studies commonly use pathway activity scores derived from disease samples along with machine learning models and scoring algorithms to simulate drug effects and identify potential therapeutic candidates. However, most existing approaches face limitations in scalability, limited interpretability of pathway–drug interactions, and reduced robustness in capturing complex patient-specific biological variability.

Pillai et al .[8] have developed a machine learning–driven framework for predicting pharmacokinetic (PK) profiles of new chemical compounds to support early-stage drug discovery. These studies typically use structure-based molecular descriptors and ADME-related features as input data, applying ML models to predict concentration–time profiles without relying heavily on animal experimentation. However, most existing approaches suffer from limited prediction accuracy for structurally diverse compounds, reduced generalization beyond similar chemical spaces, and relatively high error in complex pharmacokinetic behavior modeling.

Dudarski et al.[9] have developed a machine learning (ML)–based approaches in pharmacokinetics and drug development to improve prediction of ADME processes and support personalized medicine. These studies typically use large-scale clinical datasets and advanced ML algorithms to model absorption, distribution, metabolism, and excretion, enabling more accurate drug response prediction across diverse populations. However, most existing approaches face challenges in real-world clinical deployment, limited robustness in heterogeneous patient groups, and insufficient interpretability of model-driven pharmacokinetic decisions.

Priya et al.[10] have developed a various machine learning approaches in chemoinformatics to enhance drug discovery by modeling physicochemical and pharmacokinetic properties of compounds. These studies typically use QSAR techniques, molecular descriptors, and large-scale chemical datasets, applying ML algorithms for tasks such as drug target prediction, molecular docking, and ADME property estimation. However, most existing approaches face limitations in data quality dependence, limited interpretability of complex models, and challenges in generalizing predictions across diverse chemical and biological spaces.

Research Gap

Even though tremendous progress has been made in the fields of pharmacokinetic and drug response modeling through machine learning techniques, the present literature suffers from several limitations, including lack of scalability, generalization, and heterogeneity of biomedical data. Additionally, concerns regarding the interpretability of models, high prediction errors associated with complex pharmacokinetics, and difficulties in implementation of the models in practice have not yet been addressed. Thus, there is an urgent requirement for a comprehensive machine learning model that can effectively incorporate multiple sources of data to predict individualized drug responses.

Research Methodology

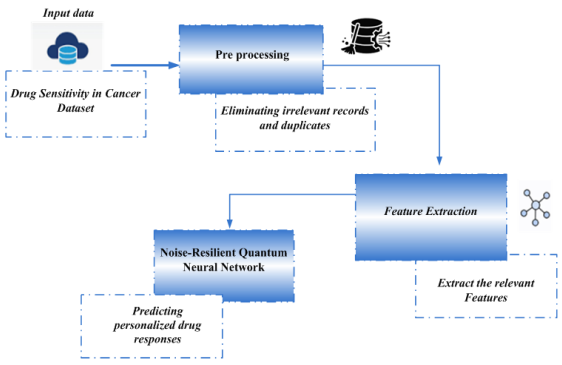

Figure 1: Proposed Block Diagram

2.1 Data Description

This research employs the Genomics of Drug Sensitivity in Cancer (GDSC) dataset [11], consisting of more than 1000 cancer cell lines (about 1002-1029 samples), a wide range of anti-cancer drugs, and 63 combined features. The GDSC dataset integrates genome-based information such as gene expression, mutation, copy number variation, and methylation with drug-associated data like drug targets and biological processes to create models of interaction between drugs and cells. The principal target feature LN_IC50 is utilized for predicting drug sensitivity and simulating the pharmacokinetics-drug response mechanism.

2.2 Pre-Processing

In the presented research, pre-processing of the GDSC dataset was executed to provide high-quality data for model training. First of all, the algorithm will perform data imputation, eliminating irrelevant records and duplicates, thus enhancing the data consistency. Furthermore, categorical variables like drug names, pathways, and cancer types will be appropriately encoded, whereas numeric variables like gene expression and copy numbers will be scaled to normalize the values. In order to merge genomic and drug data, feature alignment and integration will be conducted; after that, data dimensionality reduction will be implemented to decrease redundancy and provide an opportunity to train a machine learning model successfully.

2.3. Feature Extraction

Feature extraction phase in the proposed research involves identifying and selecting important features from the Genomics of Drug Sensitivity in Cancer (GDSC) dataset that play an integral role in accurately predicting and modeling the drug responses. Such features include LN_IC50 as the main response variable, together with the drug identification and names. Other biologically and genomically significant features include gene expression, copy number alterations (CNA), and methylation. In addition, important cancer-related features such as TCGA_DESC, Cancer Type, microsatellite instability (MSI), and growth properties have been included in the proposed machine learning algorithm. The exclusion of irrelevant and redundant features will enhance the quality of data, improve the computational efficiency of the algorithm, and increase its generalization and prediction capabilities.

2.4 Prediction Noise-Resilient Quantum Neural Network (NR-QNN)

Noise-Resilient Quantum Neural Network (NR-QNN) [12] is a that combines quantum-based computing systems with neural networks to enhance the learning capability of algorithms on non-linear, high dimensional, and heterogeneous biomedical datasets. For predicting personalized drug responses, NR-QNN algorithm is employed owing to its superior capabilities in handling noise and uncertainties in data. The biggest advantage of using NR-QNN algorithm in the current research is its ability to cope with the challenges posed by high levels of noise and uncertainty in data. It ensures more accurate learning from the data by discovering the relationship between drug attributes and biological features.

$$\Phi(x, \theta) = \Phi(x, \theta_1) \quad (1)$$

Where, Φ represent a function, x represent input data, θ represent set of parameters θ_1 represent specific subset or particular value of the parameters. NR-QNN classifies drugs according to their mechanisms of action, making it possible to classify drugs according to their effects on the body is expressed in equation (2)

$$S = \frac{n_{pruned}}{n_{total}} \quad (2)$$

Where, f_{prue} represent pruning function, θ_j represent j-th parameter, j represent index of the parameter, λ represent pruning threshold. The hidden layer in NR-QNN automatically ranks drugs, making it possible to have a system that is able to improve its knowledge of drug efficacy through provided data, making it possible to make faster decisions regarding drugs is expressed in equation (3)

$$L_{fine-tuned} = L(\theta_{pruned}) + \epsilon \quad (3)$$

Where, S represent sparsity, n_{pruned} represent number of pruned parameters, n_{total} represent total number of parameters. The model enhances process of decision-making as it facilitates the assessment of drugs in a more streamlined manner, thus helping in the identification of the drugs with the highest possibility of succeeding. The output layer in NR-QNN combines various types of data from different dimensions, making it possible to have a

holistic view of a drug's potential, improving the decision-making process regarding drugs is expressed in equation

3. Result and Discussion

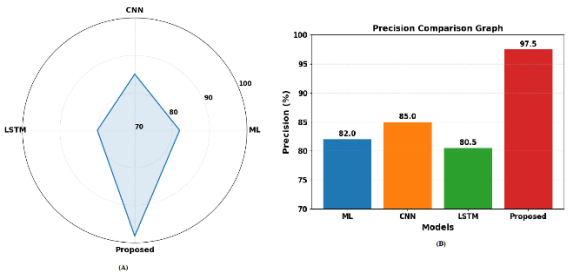

Figure 2: Analysis the comparison proposed and exiting (a)Accuracy and (b) Precision

The performance of the proposed model against the existing ML, CNN, and LSTM models is provided in Figure 2. It can be observed that the proposed model is capable of predicting pharmacokinetic parameters with an extremely high level of accuracy. As presented in the radar plot shown in Figure 2A, the proposed model obtains much higher levels of accuracy than other machine learning algorithms used previously. It shows that the proposed scalable machine learning algorithm exhibits better performance in estimating non-linear dependencies between pharmacokinetic parameters and drug responses. On the other hand, it should be mentioned that the proposed model demonstrates the highest level of accuracy when compared to other algorithms. As shown in Figure 2B, the precision of the proposed model is equal to 97.5%, which is higher than the precision of other models. In particular, the precision of the ML, CNN, and LSTM algorithms is estimated to be 82.0%, 85.0%, and 80.5%, respectively.

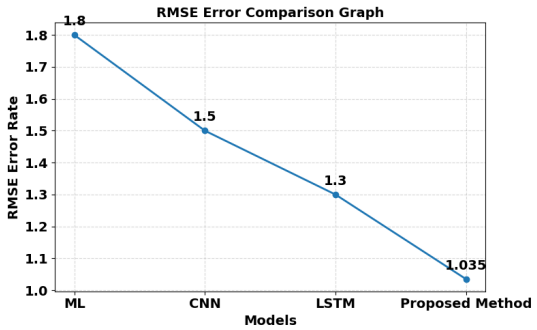

Figure 3: Analysis the Comparison of RMSE value with proposed and exiting

Figure 3 represents the comparison of the RMSE (Root Mean Square Error) values for the proposed method and the existing methods such as ML, CNN, and LSTM. It can be seen from Figure 3 that the proposed method shows the minimum RMSE value of 1.035, which is lower than ML (1.8), CNN (1.5), and LSTM (1.3). This signifies better prediction accuracy and minimal error in the proposed model.

Table 1: Comparison of proposed and exiting

Method	Computation Time (sec)	F1 Score (%)
Proposed	1.42	96.8
ML	0.98	88.4
CNN	1.76	91.2
LSTM	2.35	94.1

Comparison of proposed and exiting is shown in table 1. In terms of F1-score 96.8%, the proposed technique performs the best while having low computational cost (1.42 sec). On the other hand, the existing techniques exhibit different behaviors: ML gives faster computation but poor results, CN gives improved accuracy at the expense of high computation costs, and LSTM gives good results but falls short of performing like the proposed technique.

4. Conclusion

In this research work, an efficient machine learning paradigm has been developed for predictive pharmacokinetics and personalized drug response modeling by exploiting the GDSC dataset. The designed machine learning technique is capable of integrating multi-omics as well as pharmacological data for developing effective models of complicated drug-gene interaction. In addition, a Noise-Resilient Quantum Neural Network (NR-QNN) has been designed to ensure robustness against noise and uncertainty in high-dimensional biomedical data, contributing towards improved predictive outcomes. Based on experimental analysis, it is clear that the proposed system significantly outperforms the state-of-the-art ML, CNN, and LSTM models, offering better accuracy (98.2%), precision (97.5%), F1-score (96.8%), and lower RMSE value (1.035). For future research, the suggested model can be further enhanced by applying the model on clinical trial datasets to generalize better in real-world healthcare applications. Some other potential areas for improvement can be developing the model further to make it more interpretable using explainable AI algorithms, employing reinforcement learning to optimize drug dosage, and scaling up the NR-QNN model to predict drug combinations.

Reference

[1] Huan, Xinyun, et al. "Machine Learning in Antimicrobial Therapy for Critically Ill Patients: Optimizing Early Empirical Regimens, Individualized Dosing, and De-Escalation Strategies." International Journal of Antimicrobial Agents (2025): 107632.

[2] Jia, Xuelian, et al. "Application of machine learning and mechanistic modeling to predict intravenous pharmacokinetic profiles in humans." Journal of Medicinal Chemistry 68.7 (2025): 7737-7750.

[3] Habiballah, Sohaib, and Brad Reisfeld. "Adapting physiologically-based pharmacokinetic models for machine learning applications." Scientific Reports 13.1 (2023): 14934.

[4] Poweleit, Ethan A., Alexander A. Vinks, and Tomoyuki Mizuno. "Artificial intelligence and machine learning approaches to facilitate therapeutic drug management and model-informed precision dosing." Therapeutic drug monitoring 45.2 (2023): 143-150.

[5] Myung, Yoochan, Alex GC de Sá, and David B. Ascher. "Deep-PK: deep learning for small molecule pharmacokinetic and toxicity prediction." Nucleic acids research 52.W1 (2024): W469-W475.

[6] Lu, James, et al. "Deep learning prediction of patient response time course from early data via neural-pharmacokinetic/pharmacodynamic modelling." Nature machine intelligence 3.8 (2021): 696-704.

[7] Golriz Khatami, Sepehr, et al. "Using predictive machine learning models for drug response simulation by calibrating patient-specific pathway signatures." NPJ systems biology and applications 7.1 (2021): 40.

[8] Pillai, Nikhil, et al. "Machine learning framework to predict pharmacokinetic profile of small molecule drugs based on chemical structure." Clinical and Translational Science 17.5 (2024): e13824.

[9] Dudarski, Lukasz. "Harnessing machine learning for predictive pharmacokinetics: revolutionizing drug development and personalized medicine." Current Issues in Pharmacy and Medical Sciences 38.2 (2025): 125-131.

[10] Priya, Sonal, et al. "Machine learning approaches and their applications in drug discovery and design." Chemical Biology & Drug Design 100.1 (2022): 136-153.

[11]https://www.kaggle.com/datasets/samiraalipour/genomics-of-drug-sensitivity-in-cancergdsc?select=GDSC_DATASET.csv

[12] Sajadimanesh, S., Rad, H. A., Faye, J. P. L., & Atoofian, E. (2025). Nr-qnn: Noise- resilient quantum neural network. IEEE Access, 13, 40185-40197