Estimation of Warfarin Dosage using a Specialized XGBoost-Based Pharmacogenomic Machine Learning Model and Evaluation using XAI

Aaditya Rengarajan, Akshay Perison Davis, Navaneetha Krishnan K S, R Vishal, Subhasri Shreya S L, Jayashree L S
Department of Computer Science and Engineering
PSG College of Technology, Coimbatore, India

Abstract—Warfarin is a commonly used anticoagulant for which dosing needs to be individually optimized highly tightly to match against the potential for bleeding and thrombotic side effects. We introduce herein in this article a machine learning system that makes use of clinical, genetic, and demographic information to predict warfarin patient-specific dosing. Our method is becoming more sophisticated with various iterations starting from a baseline model, then an optimal XGBoost model incorporating polynomial feature expansions, and finally ending with an optimized gradient boosting implementation coded from scratch. Model performance is evaluated on R^2 metrics complemented with explainability tests using SHAP and LIME, hence achieving accuracy and interpretability to clinical decision-making.

Index Terms—Warfarin Dosage, XGBoost, Pharmacogenomics, Machine Learning, Explainable AI, SHAP, LIME.

I. Introduction

Warfarin continues to be one of the most commonly prescribed oral anticoagulants, yet dosing continues to be a significant problem for clinicians. The natural variability of patients, based on differences in genetic makeup, age, body mass index, dietary intake, and concomitant medications, renders it difficult to identify a "one-size-fits-all" dose. Improper dosing has dire effects, such as hemorrhagic complications in the case of overdosing or thromboembolic complications in the case of under-dosing. To this purpose, precise warfarin dosage estimation is not only a pharmacological requirement but also an essential component of patient safety and therapeutic efficacy.

Over the past few years, the science of pharmacogenomics has shed new light on how genetic diversity affects drug metabolism. For warfarin, CYP2C9 and VKORC1 genes have been found to be primarily responsible for predicting the metabolism and response of the drug. But while such genetic markers yield useful insights, clinical and demographic variables need to be incorporated to achieve an integrated model of dosage prediction. Empirical rules or linear regression techniques serve as the basis for conventional dosing algorithms, but the above techniques fail to capture the non-linear interactions between such heterogeneous variables. This study is inspired by the desire to fill the gap between data complexity that is becoming more accessible to patients and the limitations of traditional dosing approaches.

By employing sophisticated machine learning methods, i.e., gradient boosting models, we suggest a model with the ability to handle the high dimensionality and non-linearity present in pharmacogenomic data. Our solution starts with a baseline model that compares a range of traditional algorithms. This is preceded by an iterative improvement with XGBoost—a gradient boosting library of high performance well known for its performance with structured data—and augmented further by augmentation with polynomial feature transformations in order to extract non-linear trends. The final iteration of our approach introduces a locally implemented gradient boosting implementation replicating the concepts of XGBoost but written from scratch based on decision trees as base learners.

This tailored approach offers not only more control over the learning process but also has the advantage of demonstrating explicitly how each iteration enhances the overall performance of the prediction. The addition of explainable AI (XAI) techniques, i.e., SHAP and LIME, guarantees that the model's decision process is understandable. This is particularly important in a clinical environment where practitioners must know the reason behind a recommended dose of warfarin. Its contributions are two-pronged.

First, this work presents an iterative, incremental method for refining machine learning algorithms for warfarin dosing, and highlights the significance of feature engineering as well as non-linear transformations. Second, it highlights the focus on model interpretability through incorporating XAI approaches, thereby inspiring more confidence from clinical practitioners to deploy data-driven solutions. The context of this research work is, collectively, to enhance patient outcomes by minimizing the trial-and-error nature of warfarin dosing.

If clinical, genetic and demographic variables are combined in a single model, the precision of dosage prediction may be increased. This alone would be capable of minimizing the occurrence of side effects and improving the general safety of warfarin therapy. Furthermore, by producing an interpretable model, our work avoids one of the main

hindrances to implementing machine learning in medicine. This introduction not only provides the background for the ensuing technicalities, but also highlights the clinical importance that proper warfarin dosing can contribute to the management of patients. Throughout the rest of this article, we dig deeper into the literature, outline our multistep approach, report comprehensive experimental results, and end with a discussion of results and possible future work.

II. Literature Survey

The literature on warfarin dosing and pharmacogenomic modeling is vast and varied, and it crosses more than one field of study, i.e., medicine, bioinformatics, and data science. Clinical algorithms and simple linear regression models based on clinical factors such as age, weight, and concurrent medications were used in the first studies. The models, while providing a simple framework for warfarin dosing, could not capture the interactions among clinical and genetic factors. Wadelius et al.'s pioneer work showed the central role played by genetic CYP2C9 and VKORC1 polymorphisms in warfarin metabolism and hence the motivation to employ a more integrative modeling strategy [1].

Along with the augmentation of high-dimensional data, complexity in models was also augmented. Scientists started to explore machine learning approaches to improve dose prediction. Early machine learning techniques such as decision trees and support vector machines were tried on warfarin dosing with variable success. Gage et al. and Klein et al. presented some of the first data demonstrating that more sophisticated models could better estimate the optimal dose than standard linear regression models. These models were frequently condemned as being uninterpretable and hard to integrate domain expertise.[2][3]

Ensemble methods marked a significant turning point in the literature. Random Forests, where the predictions from ensembles of decision trees are aggregated, were demonstrated to be able to better identify non-linearity and interaction than those with single models. Although better is the ensemble approach such as Random Forests, they are "black-box" models and do not inform one about the way in which a specific feature adds to the overall prediction. This openness became an issue in a medical setting where knowing the reasoning behind suggesting a dosage is as important as the accuracy of the predictions.

Meanwhile, gradient boosting techniques were an available alternative. Friedman's work on gradient boosting established the theoretical foundation for techniques that iteratively improve predictions into being more accurate by making incremental improvements on residual errors. XGBoost, a particular variant of gradient boosting, was made popular because it was scalable, performance-oriented, and superior on structured data.[4] Various studies proved that XGBoost could well capture non-linear interactions between clinical and genetic variables in warfarin dosing. For example, one study by Asiimwe et al. compared more than twenty machine learning algorithms and established that ensemble techniques, specifically gradient boosting-based ones, achieved a significant predictive performance gain. [10]

Also, research using feature engineering methods like polynomial expansion has demonstrated that adding interaction terms can further increase the capacity of the model to learn sophisticated relationships. In one study, Huang et al. used polynomial transformations to clinical data and found an unprecedented improvement in model performance. [6]While these methods improve the accuracy of the model, they also enhance the model's complexity, for which strong techniques of model interpretability are required.

The problem of interpretability has generated rising interest in explainable AI (XAI) techniques. A couple of the more popular approaches of XAI are SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations). SHAP, a game-theoretic method, is a global view of feature contributions with each feature receiving a value of contribution in the prediction. LIME, meanwhile, obtains local interpretability by approximating the model near a single prediction with an interpretable linear model. Various tests have compared the two methods, and most of them have concluded that although LIME can be useful for making predictions, SHAP provides a more overall and consistent explanation of what the model is doing. The integration of these methods in warfarin dose models is particularly beneficial, as clinicians need clarity to have faith in the AI recommendations.

Additional studies by Lee et al. investigated deep learning methods like recurrent neural networks (RNNs) and long short-term memory (LSTM) networks for warfarin dosing.[11] While highly accurate, such models were too complex and computationally demanding for clinical applications. Lower-accuracy but more interpretable models were employed in clinical trials instead. This tension between accuracy and interpretability is seen repeatedly throughout the literature.

Recent research also considered incorporating the lifestyle and socioeconomic aspects into dose models. One such model has been proposed by Anand et al. where, although cheaper, it ignored the pharmacogenetic markers with resulting lower prediction capacity.[5] On the other hand, those studies including genetic and

clinical information, i.e., like by Mohamed et al. and Liu et al., have achieved a better capacity through exploiting the complementarity of both types of information.[7][8]]

Overall, the literature emphasizes the progression from straightforward linear models to sophisticated ensemble and gradient boosting techniques, alongside a concurrent focus on model interpretability. Although the initial work was pioneering through the identification of the principal genetic markers, later work has involved the addition of extra variables as well as more advanced machine learning algorithms. The employment of XGBoost, specifically, has been in widespread practice due to the fact that it can handle non-linear interaction and scalability. However, the question lies in how one can achieve very high prediction accuracy without compromising model interpretability—a trade-off that is critically significant in a clinical setting. This literature review not only points to the progress in warfarin dosing models but also to the weaknesses that our contribution will plug, namely through iterative model refinement and explainable AI incorporation.

III. Methodology

This is the full description of the methodology that was used in deriving a prediction model to estimate warfarin dosage. Step-by-step, our method will be to refine and enhance the model performance progressively. Methodology has various phases, such as data preprocessing, feature transformation and engineering, baseline model, and iterative refinement. Further, we explain the gradient boosting implementation to closely resemble XGBoost but defined manually to aid in better understanding how the model learns.

A. Data Cleaning and Preprocessing

Data preprocessing is a very important part of any machine learning pipeline. Data for this research involves patient history with clinical values, demographics, and genetic profiles. Preprocessing has multiple stages. First, missing values in the data set are addressed by applying both KNN imputation for numerical attributes and mode imputation for categorical attributes. For a high missing rate, median imputation is applied to prevent the incorporation of outlier values into the distribution. The step is important in ensuring the integrity of the data set and preventing the model from learning from incomplete data. Then, normalization of numerical features is done through min-max scaling. This is to ensure that all features are equally contributing to the learning process and that the model does not bias towards features with larger numerical ranges. One-hot encoding is then applied to the categorical features like gender and ethnicity. One-hot encoding converts categorical features into a matrix of binary features, wherein the model can handle categorical features numerically. Outliers are identified using the Interquartile Range (IQR) technique. Outliers based on data entry errors or rare patient statuses are eliminated so that they won't perform negatively on the model. Except for the above methods, missing genetic information for genes like VKORC1 and CYP2C9 are imputed using the distribution observed in the training set in a way that variance is preserved.
Feature engineering is required in an effort to capture the high-level interactions within pharmacogenomic data. We utilize both domain knowledge and automated feature construction methods in our methodology. We begin by finding the Body Mass Index (BMI)—from the given values of height and weight. The BMI is obtained as the ratio of weight to height squared. The transformation is due to the fact that BMI is known to impact drug distribution and metabolism. Then, we utilize polynomial feature transformation to pick up non-linear relations among continuous features. For instance, based on height and weight features, the transformation includes other features: height^2, weight^2, and the product of height and weight. The justification for polynomial expansion is so that the model is able to learn high-order interactions that otherwise would be lost when learning linear relations alone. This phase significantly increases the dimensionality of the feature space, which, being computationally more expensive, enables the model to learn more nuanced patterns of the data.

B. Baseline Model Building using XGBoost

The initial major implementation of our approach uses the XGBoost algorithm. XGBoost, or eXtreme Gradient Boosting, is a very fast and scalable implementation of gradient boosting that has proven to have extremely good performance on structured data. Transformed and preprocessed data are used in an XGBoost regression model in this implementation. The model is learned to reduce the mean squared error (MSE) between actual and predicted dosages of warfarin. Hyperparameters at training are tree depth, learning rate, and subsample ratio, which are optimized to work best. The initial XGBoost model achieved an R^2 score of 46.1%, serving as a benchmark for further improvements.

C. Improved Model with Polynomial Feature Transformation

The second iteration adds polynomial feature transformations to the basic XGBoost model. Already covered, adding higher-order features allows the model to learn about non-linear relationships among predictors. As part of this iterative optimization, the XGBoost model is retrained on the expanded set of features. All the boosting trees in the ensemble can now utilize the polynomial features and the original features as well, gaining a better understanding of data structure. Including those transformations had a tremendous performance boost, since the model achieved an R^2 of 76.13%.

D. Custom Gradient Boosting Implementation: Detailed Explanation

The last and most complete form of our solution is the development of a custom gradient boosting algorithm. The custom implementation is designed to mimic the essential ideas of XGBoost but with greater transparency and learning process control. The custom gradient boosting model is obtained through the use of decision trees as weak learners. It starts by estimating an initial prediction that is simply the mean of the target variable (warfarin dose) over the training data.

1) First Prediction and Residual Computation: The algorithm begins with the prediction of the mean warfarin dose for every patient. The residual error for each case is calculated as a difference between the initial prediction and this first prediction. This residual forms the foundation for the first weak learner.

2) Training the First Decision Tree: A decision tree regressor is trained on the residuals. The aim of this tree is to capture patterns in the error distribution not picked up by the initial model. The tree is set up with a maximum depth (say 8) to avoid overfitting and keep the model generalizable.

3) Incremental Update using a Learning Rate: After the decision tree is trained, its predicted values are scaled down by a small learning rate (e.g., 0.05) and incorporated into the initial predictions. The use of a small learning rate prevents the model from immediately making gigantic prediction adjustments and thereby avoids overfitting.

4) Iterative Residual Learning: The model calculates updated residuals based on updated predictions. The residuals are then used to learn a new decision tree and then combine its scaled prediction with the present prediction. This procedure is repeated a predetermined number of times, and in each step, it decreases the error further by targeting the residual remaining. Early stopping is applied to stop training if mean squared error decrease plates for some iterations so that overfitting can be avoided.

This specialized implementation not only replicates the iterative process of error correction of XGBoost but also provides fine-grained ability to tune every hyperparameter. Decision trees, or weak classifiers, are trained in an iterative manner to give the most importance to the best residual errors. The collective result of these incremental adjustments is a model that reliably estimates the intricate interaction of input features and warfarin dosage.

In short, our approach ranges from extensive data preprocessing and intricate feature engineering to a cutting-edge ensemble learning architecture. The transformation of a simple XGBoost model into a highly sophisticated custom-built gradient boosting model captures the spirit of model fine-tuning in iterations. After incorporating polynomial features and taking advantage of the power of sequential learning, the ultimate model yields an R^2 value of 94.36%, much greater than in predictive accuracy. The transparency of the custom implementation also makes the model more suitable for clinical applications, where usability is just as important as performance.

5) Model Checking and Loss Curve Plotting: During the training of the model, the performance of the model is checked based on the mean squared error (MSE). It is plotted with a loss curve to see how MSE goes down in later iterations. Based on this curve, the trend of convergence of the algorithm can be studied and the optimal number of boosting iterations determined. The pseudocode for the custom gradient boosting algorithm is described in Algorithm 1.

Algorithm 1: Algorithm for Custom Gradient Boosting

```
class CustomGradientBoosting:
    def __init__(self, n_estimators=100,
        learning_rate=0.05, max_depth=8):
        self.n_estimators = n_estimators
        self.learning_rate = learning_rate
        self.max_depth = max_depth
        self.models = []
        self.loss_history = []

    def fit(self, X, y):
        self.init_pred = np.mean(y)
        y_pred = np.full(y.shape,
            self.init_pred)
        for i in range(self.n_estimators):
            residuals = y - y_pred
            tree = DecisionTreeRegressor(\
                    max_depth=self.max_depth)
            tree.fit(X, residuals)
            self.models.append(tree)
            y_pred += self.learning_rate *
                tree.predict(X)
            mse = mean_squared_error(y,
                y_pred)
            self.loss_history.append(mse)
            if i > 20 and mse >
                self.loss_history[-20]:
                break

    def predict(self, X):
        y_pred = np.full((X.shape[0],),
            self.init_pred)
        for tree in self.models:
            y_pred += self.learning_rate *
                tree.predict(X)
        return y_pred
```

IV. Results

The experimental comparison of the warfarin dosage prediction models shows a consistent improvement in performance with each iteration. The first XGBoost model, which was trained on the preprocessed data without polynomial feature expansion, had an R^2 score of 46.1%. While the model was a good starting point, it was evident that the linear assumptions used in the baseline approach were not enough to capture the non-linear interactions in the data.

With the addition of polynomial feature transformations, the XGBoost model's performance significantly enhanced to an R^2 of 76.13%. The reason for this was that the improved capacity of the model to identify interaction effects among variables like height, weight, and BMI. The introduction of higher-order terms enabled the algorithm to identify intricate relationships that are of the utmost importance in warfarin's pharmacokinetics.

Metric	Value
Mean Squared Error (MSE)	16.177956176667514
R-squared (R^2)	0.94353946983268
R^2 Score (%)	94.36%

Table I: Final Iteration Model Evaluation Metrics

The final iteration, the custom gradient boosting, yielded an R^2 of 94.36% (ref Fig. 1). This round is a radical improvement in performance because the custom model was able to iteratively refine residual errors to a very high degree of accuracy. A closer examination of the loss curve reveals a consistent reduction in mean squared error with boosting iterations, which means that the model was learning well from the residuals. Early stopping criteria were used so that training was terminated before the model started to overfit the training set.

Other performance measures like Root Mean Squared Error (RMSE) were tracked in addition to R^2 measure during training. RMSE reduced incrementally, demonstrating how the model's accuracy at predicting warfarin dose rises. Loss curve plots indicate that the model consistently converges and additional iterations yield decreasing returns.

Feature	Value	Effect on Prediction
Gender_male	1	↑ (increase)
Target INR	2.5	↑ (increase)
Amiodarone (Cordarone)	0	↑ (increase)
Height (cm)	2.427	↑ (increase)
Weight (kg)	1.7	↑ (increase)
Aspirin	1	↓ (decrease)
BMI	31.05	↓ (decrease)
Age	0.08328	↓ (decrease)

Table II: Feature Contributions to Model Prediction

Feature	Effect	Reason
Cyp2C9 Genotypes (*1/*2 & *1/*3) = 0	Increases	No genetic mutation → Warfarin metabolized faster
Gender = Male	Increases	Men need higher doses than women
Target INR = 2.5	Increases	Higher INR target → needs stronger effect
Amiodarone (Cordarone) = 0	Increases	Not taking Amiodarone → No drug interaction reducing dose
Height & Weight	Increases	Larger body → Needs more Warfarin
Aspirin = 1	Decreases	Aspirin is also a blood thinner → Reduces Warfarin requirement
BMI = 31.05	Decreases	More fat storage → Less active Warfarin in blood
Age (Older)	Decreases	Older people metabolize Warfarin slower
INR on Reported Dose	Decreases	Higher INR → Patient is sensitive to Warfarin

Table III: Feature Contribution from SHAP

Explainability was of key interest in our assessment. SHAP and LIME were hence utilized for explanation of prediction. SHAP global explanation provided every one of the features' contribution to the resultant prediction as a numerical value (ref. Fig.2). Such a feature as CYP2C9 genotype, VKORC1 status, age, BMI, and concomitant medication presence were found to contribute significantly

positively or negatively to the predicted dose(ref. Table I). The SHAP value distributions provided insight into the way these features interact within the model framework, thus enhancing the clinical validity of the predictions.

Table IV: LIME Feature Importance for Dose Prediction

Feature	LIME Value (Mean)	Effect on Dose
CYP2C9 Genotype (*1/*3)	-5.7	Decreases dose
VKORC1 Genotype (A/A)	-4.2	Decreases dose
Age (75 years)	-2.8	Decreases dose
BMI (30.2)	+3.6	Increases dose
Weight (85 kg)	+3.9	Increases dose
Amiodarone Use	-6.1	Decreases dose
Aspirin Use	-1.2	Decreases dose
Target INR (2.5)	+4.4	Increases dose

LIME was utilized to generate localized explanations for point predictions. Through the approximation of the behavior of the complex model around a particular data point by an interpretable linear model, LIME enabled us to check that the model's decision-making process conformed to clinical expectations (ref. Fig. 3.1). For instance, in scenarios where a patient had a known genetic variation that linked the patient to slower metabolism, LIME emphasized the relative negative weight in the dosage prediction (ref. Fig. 3.2).

The use of quantitative performance metrics and qualitative interpretability tests together highlights the robustness of the final model. The outstanding R^2 value of 94.36% clearly shows that the tailored gradient boosting model can explain almost all warfarin dosing variability. Moreover, the use of XAI techniques ensures the model predictions not only accurate but also transparent and reliable for clinical decision-making.

Overall, the experimental outcomes confirm the effectiveness of our iterative process. The journey from a vanilla XGBoost model to a polynomial-featured model, and then to an in-house gradient-boosting implementation, testifies to the significance of both feature engineering and algorithmic design. The outcomes also prove that an well-engineered machine learning model can mitigate the inherent complexity of warfarin dosing while yielding understandable and actionable insights.

V. Conclusion

We introduce in this paper the complete methodology of warfarin dosage prediction using a tailored XGBoost-based pharmacogenomic machine learning algorithm, culminating in a tailored gradient boosting implementation. Our methodology progressed in various iterative stages from simple baseline models to improved feature transformations to an entirely tailored implementation. The final implemented model had an R^2 of 94.36%, which is notably improved predictive performance compared to the initial methods. Furthermore, the incorporation of explainable AI methods, including SHAP and LIME, provided insightful information regarding the feature contributions, thereby enhancing the model interpretability for doctors.

The findings that are described here highlight the possibility of synergy between cutting-edge machine learning methods and strong data preprocessing and feature engineering to tackle the subtle challenges of warfarin dosing. Along with revealing a correct high-performing prediction model, our contribution provides state-of-the-art computational methods into the clinical practice setting. The rigorous interpretability analysis also allows clinicians to have confidence in the model recommendations, which is of paramount importance to its use in actual clinical practice.

Future research will involve prospectively validating the model in clinical trials and investigating its integration with electronic health records (EHR) systems. Actions like these are needed to further calibrate the model and to ensure that it can be easily integrated into clinical practice. Overall, the results of this study are an encouraging step towards safer and more personalized warfarin therapy and illustrate the revolutionary potential of explainable machine learning to revolutionize healthcare.

References

[1] Wadelius et al. Wadelius, M., Chen, L. Y., Eriksson, N., Bumpstead, S., Ghori, J., Wadelius, C., Bentley, D., McGinnis, R., & Deloukas, P. (2007). Association of warfarin dose with genes involved in its action and metabolism. Human Genetics,

[2] Gage et al. Gage, B. F., Eby, C., Johnson, J. A., Deych, E., Rieder, M. J., Ridker, P. M., Milligan, P. E., Grice, G., Lenzini, P., Rettie, A. E., & Aquilante, C. L. (2008). Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. Clinical Pharmacology & Therapeutics

[3] Klein et al. Klein, T. E., Altman, R. B., Eriksson, N., Gage, B. F., Kimmel, S. E., Lee, M. T. M., Limdi, N. A., Page, D., Roden, D. M., Wagner, M. J., Caldwell, M. D., & Johnson, J. A. (2009). Estimation of the warfarin dose with clinical and pharmacogenetic data. New England Journal of Medicine

[4] Friedman on Gradient Boosting Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5), 1189–1232.

[5] Anand et al. Anand, S. S., Yusuf, S., Xie, C. C., Pogue, J., Eikelboom, J., & Mehta, S. R. (2007). Oral anticoagulation in patients with coronary artery disease: Insights from the OASIS trials. Archives of Internal Medicine

[6] Huang et al., "Nonlinear ML Models for Warfarin Dose Prediction," IEEE Transactions on Biomedical Engineering, Year.

[7] Mohamed et al., "Comparative Analysis of Warfarin Dosing Models in Latin American Populations," Pharmacogenomics Journal, Year.

[8] Liu et al., "Ensemble Learning for Improved Warfarin Dose Prediction," IEEE Access, Year.

[9] Gu et al., "Adapted Neural-Fuzzy Inference Systems for Warfarin Dose Estimation," Computers in Biology and Medicine, Year.

[10] Asiimwe et al., "Evaluation of 21 Machine Learning Models for Warfarin Dose Prediction," Journal of Medical Systems, Year.

[11] Lee et al., "Deep Learning for INR Prediction and Warfarin Dosing," IEEE Journal of Biomedical and Health Informatics, Year.