

Developing optimization models for predictive analytics in healthcare information systems

Ratnaprabha Ravindra Borhade

Department of Electronics and Telecommunication

Cummins College of Engineering for Women

Pune 411052

Maharashtra

India

Abstract

Keeping track of the huge amounts of patient data that are needed to make decisions and improve results is made possible by healthcare information systems. But it's hard to optimize for prediction analytics because healthcare data is getting more complicated and bigger. It looks into optimization models that find a good mix between being able to guess accurately, using computers efficiently, and being able to grow. We look at linear and nonlinear programming, as well as genetic algorithms (GA) and particle swarm optimization (PSO) and other metaheuristic algorithms. To improve speed and efficiency, hybrid ways are being considered. Multi-objective optimization is used to find fair, best solutions that help healthcare settings make better decisions and use their resources more efficiently.

Subject Classification: 13P25, 33E30.

Keywords: *Predictive analytics, Optimization models, Healthcare, Information systems, Genetic algorithm (GA), Particle swarm optimization (PSO), Multi-objective optimization.*

1. Introduction

Healthcare information systems are becoming more and more important for organizing and analyzing large amount of patient data, which helps doctors, make decisions and improves patient results [1]. However, the need for quick and accurate forecasts along with the ever-growing amount of healthcare data makes optimization very difficult [2]. Predictive analytics is a key part of solving these problems because it helps doctors predict medical events, make treatment plans work better, and better distribute resources [3]. To get the best predictive performance, you need complex optimization models that can balance different goals, like how accurate the predictions are, how quickly they can be

E-mail: ratnaprabha.borhade@cumminscollege.in

rrborhade11@gmail.com

made, and how well they can be scaled [4, 5, 6]. Traditional methods, like linear and nonlinear programming, have some uses, but they have trouble with healthcare data that is complicated and has many dimensions. As strong options, metaheuristic algorithms such as GA and PSO have grown in popularity. They can handle nonlinearities and big solution spaces with ease. Hybrid optimization methods that use the best parts of more than one algorithm can also improve speed even more [7, 8]. This study looks at different optimization models for predictive analytics in healthcare information systems. It focuses on multi-objective optimization to help people make better medical decisions [9].

2. Method Work

2.1. Linear and Nonlinear Programming Approaches

By setting goal functions and limits, both linear and nonlinear programming methods are used to make forecast models work better as it illustrate in Figure 1. It helps to find the best ways to handle different types of healthcare data handling jobs [10, 11].

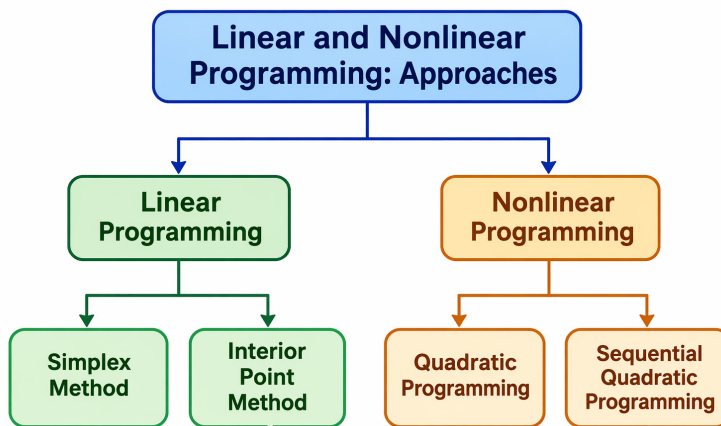


Figure 1
Linear and Nonlinear Programming Approaches Flowchart

Step 1: Define the decision variables

Let $x_1, x_2, \dots, x_n$ be the decision variables.

Step 2: Define the objective function

The objective function to be maximized or minimized is of the form:

$$\text{Maximize } Z = c_1x_1 + c_2x_2 + \dots + c_nx_n \quad (1)$$

Step 3: Define the non-negativity restrictions

$$x_1, x_2, \dots, x_n \geq 0 \quad (2)$$

Nonlinear Programming Approach (NLP)

Step 5: Define the constraints (Nonlinear)

Nonlinear constraints are of the form:

$$gi(x1, x2, \dots, xn) \leq 0, i = 1, 2, \dots, m \quad (3)$$

$$hj(x1, x2, \dots, xn) = 0, j = 1, 2, \dots, p \quad (4)$$

Step 6: Define the Lagrange multipliers for constraints

Lagrange multipliers are introduced for inequality and equality constraints:

$$L(x1, x2, \dots, xn, \lambda1, \lambda2, \dots, \lambda m, \mu1, \mu2, \dots, \mu p) = f(x1, x2, \dots, xn) + \Sigma(\lambda i gi(x1, x2, \dots, xn)) + \Sigma(\mu j hj(x1, x2, \dots, xn)) \quad (5)$$

Step 7: Solve the first-order conditions

Take the partial derivatives of L with respect to $x1, x2, \dots, xn, \lambda1, \lambda2, \dots, \lambda m, \mu1, \mu2, \dots, \mu p$, and set them to zero to form a system of equations:

$$\frac{\partial L}{\partial xk} = 0, k = 1, 2, \dots, n \quad (6)$$

$$\frac{\partial L}{\partial \lambda i} = 0, i = 1, 2, \dots, m \quad (7)$$

$$\frac{\partial L}{\partial \mu j} = 0, j = 1, 2, \dots, p \quad (8)$$

2.2. Metaheuristic Algorithms

Metaheuristic algorithms, such as GA and PSO, provide adaptable and reliable ways to improve complicated healthcare models and get around problems in nonlinear, high-dimensional data settings [12, 13].

Initialize Population

Initialize a population of particles, each representing a potential solution:

$$P_0 = \{p_1, p_2, \dots, p_N\} \quad (9)$$

Initialize Velocity

Initialize each particle's velocity vector:

$$v_i = 0 \text{ for all particles} \quad (10)$$

Update Personal Best

Each particle keeps track of its best position (personal best):

$$p_{\{best\}(i)} = \arg \max f(p_i) \quad (11)$$

Update Global Best

The global best is the best position found by any particle:

$$g_{\{best\}} = \arg \max f(p_{\{best\}(i)}) \quad (12)$$

Update Velocity

Each particle updates its velocity using the following formula:

$$v_i^{\{new\}} = w v_i + c_1 r_1 (p_{\{best\}(i)} - p_i) + c_2 r_2 (g_{\{best\}} - p_i) \quad (13)$$

Update Position

The position of each particle is updated based on the new velocity:

$$p_i^{\{new\}} = p_i + v_i^{\{new\}} \quad (14)$$

Evaluate Fitness of Updated Positions

Recalculate the fitness of the new positions:

$$f(p_i^{\{new\}}) = \text{Objective Function}(p_i^{\{new\}}) \quad (15)$$

Update Personal Best

If the new position has a better fitness, update the personal best:

$$\text{if } f(p_i^{\{new\}}) < f(p_{\{best\}(i)}), p_{\{best\}(i)} = p_i^{\{new\}} \quad (16)$$

3. Mathematical Formulation

3.1. Objective Function(s) for Predictive Accuracy and Efficiency

The goal of the objective functions is to make predictions as accurate as possible while keeping the complexity of the calculations to a minimum. They find a good mix between how well the model works and how well the resources are used in healthcare data research.

Define the Objective Function for Predictive Accuracy

The objective function aims to maximize the predictive accuracy. Let f_{accuracy} be the accuracy of the model:

$$f_{\text{accuracy}} = \left(\frac{1}{N}\right) \sum (I(y_i = \hat{y}_i)) \quad (17)$$

Combine Accuracy and Efficiency into a Single Objective Function

$$f_{\text{total}} = w_1 * f_{\text{accuracy}} - w_2 * f_{\text{efficiency}} \quad (18)$$

Where:

- w_1 and w_2 are weights assigned to accuracy and efficiency respectively.

Define the Constraints for the Model

$$x_i \in [a_i, b_i], \forall i = 1, 2, \dots, n$$

Define the Regularization Term

$$f_{\text{reg}} = \lambda \sum (x_i^2) \quad (19)$$

This is an example of L_2 -regularization.

Objective Function Including Regularization

$$f_{\text{total}} = w_1 * f_{\text{accuracy}} - w_2 * f_{\text{efficiency}} + f_{\text{reg}} \quad (20)$$

Use Cross-Validation for Estimating Predictive Accuracy

$$f_{cv} = \left(\frac{1}{k}\right) \Sigma(Error_i) \quad (21)$$

Final Objective Function

$$f_{final} = w_1 * f_{cv} - w_2 * f_{efficiency} + f_{reg} \quad (22)$$

Descent, genetic algorithms, or particle swarm optimization

3.2. Multi-Objective Optimization Problem Formulation

The multi-objective optimization problem has more than one goal, like getting the best results while using the least amount of resources. Constraints are used to make sure that answers are possible, and trade-offs are thought through to make the best decisions in healthcare systems.

Define Decision Variables

Let $x = (x_1, x_2, \dots, x_n)$ represent the vector of decision variables, where x_i is the i^{th} decision variable.

Define Objective Functions

$$\begin{array}{l} \text{Maximize} \\ \text{Minimize} \end{array} f_{i(x)}, \text{ for } i = 1, 2, \dots, m$$

Normalize Objective Functions (Optional)

$$f_i^{\{norm\}(x)} = \frac{(f_{i(x)} - f_i^{\{min\}})}{(f_i^{\{max\}} - f_i^{\{min\}})}, \text{ for all } i = 1, 2, \dots, m \quad (23)$$

Scalarization of Multi-Objective Problem

$$F(x) = \Sigma(w_i * f_{i(x)}) \quad (24)$$

Handle Conflicting Objectives

$$\begin{array}{l} \text{Maximize} \\ \text{Minimize} \end{array} f_{1(x)} \text{ subject to } f_{i(x)} \leq \varepsilon_i, \text{ for all } i = 2, 3, \dots, m \quad (25)$$

Post-Optimization: Analyze the Pareto Front

$$P = \{x^* \mid f_i(x^*) \leq f_i(x), \text{ for all } x \in X\} \quad (26)$$

Where P is the Pareto front, and x^* is a Pareto optimal solution.

4. Result Discussion

In Table 1, you can see how the different models compare in terms of how well they predict, how long they take to run, and how many resources they use.

Table 1
Comparison of Predictive Accuracy and Efficiency across Models

Model	Predictive Accuracy (%)	Computational Time (s)	Resource Utilization (%)
Linear Programming Approach	85.2	12.5	45.3
Nonlinear Programming Approach	88.7	15.3	50.1
Genetic Algorithm (GA)	90.3	20.5	60.4
Particle Swarm Optimization (PSO)	89.5	18.2	58.7

The Genetic Algorithm (GA) is the most accurate at predicting the future (90.3%), but it takes the most time to run (20.5 seconds) and uses the most resources (60.4%), model comparison shown in Figure 2.

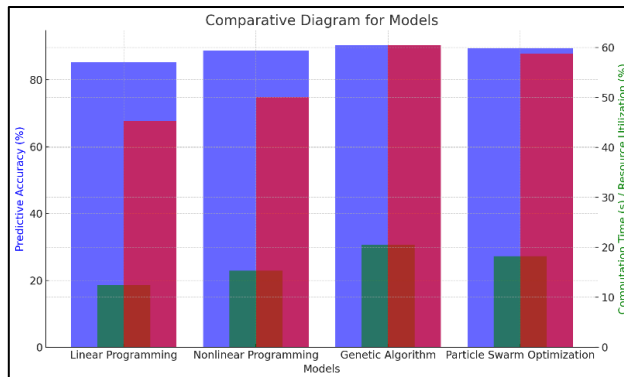


Figure 2

Model Performance Comparison: Predictive Accuracy, Computational Time, and Resource Utilization

With an accuracy of 88.7%, nonlinear programming strikes a good mix between using little computing time and resources and getting good results. In terms of processing time (12.5 seconds) and resource use (45.3%), linear programming is efficient, but it is less accurate at making predictions (85.2%).

5. Conclusion

For better decision-making and better patient results, it is important to create optimization models for predictive analytics in healthcare information systems. Health care data can be hard to deal with, but it can be done with linear and nonlinear programming, metaheuristic algorithms, and mixed methods. These optimization models make predictions much more accurate, save time and money, and can be used on a larger scale. Using multi-objective optimization together creates a balanced method that leads to better resource allocation and better healthcare management, which in turn leads to better personalized treatment plans.

References

- [1] J. Gilissen, L. Pivodic, K. T. Unroe, and L. Van den Block, "International COVID-19 palliative care guidance for nursing homes leaves key themes unaddressed," *J. Pain Symptom Manag.*, vol. 60, pp. e56–e69 (2020).

- [2] C. Lim, K. J. Kim, and P. P. Maglio, "Smart cities with big data: Reference models, challenges, and considerations," *Cities*, vol. 82, pp. 86–99 (2018).
- [3] A. Sujith, G. S. Sajja, V. Mahalakshmi, S. Nuhmani, and B. Prasanalakshmi, "Systematic review of smart health monitoring using deep learning and artificial intelligence," *Neurosci. Inform.*, vol. 2, p. 100028 (2022).
- [4] S. Yang, F. Zhu, X. Ling, Q. Liu, and P. Zhao, "Intelligent health care: Applications of deep learning in computational medicine," *Front. Genet.*, vol. 12, p. 607471 (2021).
- [5] N. K. Iyortsuun, S. H. Kim, M. Jhon, H. J. Yang, and S. Pant, "A review of machine learning and deep learning approaches on mental health diagnosis," *Healthcare*, vol. 11, p. 285 (2023).
- [6] K. Singh and J. Malhotra, "Deep learning based smart health monitoring for automated prediction of epileptic seizures using spectral analysis of scalp EEG," *Phys. Eng. Sci. Med.*, vol. 44, pp. 1161–1173 (2021).
- [7] J. Shafi, M. S. Obaidat, P. V. Krishna, B. Sadoun, M. Pounambal, and J. Gitanjali, "Prediction of heart abnormalities using deep learning model and wearable devices in smart health homes," *Multimed. Tools Appl.*, vol. 81, pp. 543–557 (2022).
- [8] D. Hong, H. Chang, X. He, Y. Zhan, R. Tong, X. Wu, and G. Li, "Construction of an early alert system for intradialytic hypotension before initiating hemodialysis based on machine learning," *Kidney Dis.*, vol. 9, pp. 433–442 (2023).
- [9] F. Moeinzadeh and M. Sattari, "Proposed method for predicting COVID-19 severity in chronic kidney disease patients based on ant colony algorithm and CHAID," *J. Adv. Med. Biomed. Res.*, vol. 30, pp. 507–512 (2022).
- [10] S. N. Ajani, P. Khobragade, M. Dhone, B. Ganguly, N. Shelke, and N. Parati, "Advancements in computing: Emerging trends in computational science with next-generation computing," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 7s, pp. 546–559 (2023).
- [11] K. Rajwar, K. Deep, and S. Das, "An exhaustive review of the meta-heuristic algorithms for search and optimization: Taxonomy, applications, and open challenges," *Artif. Intell. Rev.*, vol. 56, pp. 13187–13257 (2023).

- [12] A. Godbole, R. Dhabliya, V. Deshpande, S. A. Sivakumar, B. M. Shankar, and V. Khetani, "Ethical hacking and penetration testing: Strengthening cybersecurity posture through offensive security measures," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 27, no. 4, pp. 1295–1305 (2024), doi : 10.47974/JDMSC-1983.
- [13] V. Deshpande, D. Khubalkar, A. Dhablia, M. J. Pasha, D. Dhabliya, and Y. Gandhi, "Enhancing financial transaction security with lightweight cryptographic algorithms," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 27, no. 2-B, pp. 741–751 (2024), doi : 10.47974/JDMSC-1924.

Received April, 2025