العدد 15 تشرين2 2024
No.15 Nov 2024
المجلة العراقية للبحوث الإنسانية والإجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN 2790-1254

Optimizing Machine Learning Models for Predictive Analytics in Healthcare Using Advanced Numerical Methods

Saja Mahdi AL-Khafaji
Najaf Education Directorate,Ministry of Education,Repuublic of Iraq
Email: sajamahdi991@gmail.com
Noor Hasan Kadhim
Diwaniyah Education Directorate,Ministry of Education,Repuublic of Iraq

Abstract

This research paper aims to study the use of higher accuracy numerical implementation techniques in the enhancement of the rates of machine learning models for predictive analysis in the health sector. The study responds to the issues linked with huge and composite datasets – the general attributes of most healthcare datasets. Through the implementation of numerous computational optimization methods as gradient descent and Newton’s method, the study intends to improve the efficiency of several machine learning algorithms like neural networks and decision tree. The approach that can be described as a logical flow of steps includes collection of data, preprocessing of data and numerical computations within the scope of the machine learning paradigm. A case study involving prediction of readmission rates of patients put into the appropriate model shows how these integrated techniques can be employed and the influence of each of them. These outcomes show promising effects regarding model performance and computational cost, which prove the necessity of utilizing sophisticated number theory concepts in the healthcare analysis. Thus, this work can be regarded as the continuation of the AI-focused developments in an attempt to improve patient’s wellbeing and streamline processes in the sphere of healthcare.

Keywords: Artificial Intelligence; Gradient Descent; Healthcare Analytics; Machine Learning; Numerical Methods; Predictive Analytics; Patient Readmission; Newton's Method; Data Preprocessing

تحسين نماذج التعلم الالي للتحليلات التنبؤية في الرعاية الصحية بأستخدام الأساليب العددية المتقدمة

م.م. سجى مهدي جليل الخفاجي
نور حسن كاظم

الخلاصة

تهدف هذه الورقة البحثية إلى دراسة استخدام تقنيات التنفيذ العددي عالية الدقة في تعزيز معدلات نماذج التعلم الآلي للتحليل التنبؤي في القطاع الصحي. تستجيب الدراسة للقضايا المرتبطة بمجموعات البيانات الضخمة والمركبة - السمات العامة لمعظم مجموعات بيانات الرعاية الصحية. من خلال تنفيذ العديد من أساليب التحسين الحسابية مثل النسب المتدرج وطريقة نيوتن، تهدف الدراسة إلى تحسين كفاءة العديد من خوارزميات التعلم الآلي مثل الشبكات العصبية وشجرة القرار. يتضمن النهج الذي يمكن وصفه بالتدفق المنطقي للخطوات جمع

البيانات والمعالجة المسبقة للبيانات والحسابات الرقمية ضمن نطاق نموذج التعلم الآلي. توضح دراسة الحالة التي تتضمن التنبؤ بمعدلات إعادة قبول المرضى الموضوعة في النموذج المناسب كيف يمكن استخدام هذه التقنيات المتكاملة وتأثير كل منها. تظهر هذه النتائج تأثيرات واعدة فيما يتعلق بأداء النموذج والتكلفة الحسابية، مما يثبت ضرورة استخدام مفاهيم نظرية الأعداد المتطورة في تحليل الرعاية الصحية. وبالتالي، يمكن اعتبار هذا العمل بمثابة استمرار للتطورات التي تركز على الذكاء الاصطناعي في محاولة لتحسين رفاهية المريض وتبسيط العمليات في مجال الرعاية الصحية.
الكلمات المفتاحية: الذكاء الاصطناعي؛ نزول متدرج؛ تحليلات الرعاية الصحية؛ التعلم الالي؛ الطرق العددية؛ التحليلات التنبؤية؛ إعادة قبول المريض؛ طريقة نيوتن؛ المعالجة المسبقة للبيانات.

NTRODUCTION

Clinical prediction systems in the health sector is a predictive analytics tool that is built to forecast events or trends likely to be encountered in patient treatment in order to enhance decision making regarding the institution's operations. With the help of big data and statistical analysis, prognosis tools can help in areas like the emergence of diseases, patients' readmissions, and treatment outcomes (Abdelaziz et al. , 2018; Golas et al. , 2018). The introduction of big data facts and options as well as ML technologies has greater the healthcare facts evaluation and prediction capacity (Fei et al. , 2021; Ngiam & Khor, 2019).

However, there are still some issues to tackle in improving predictive models' high accuracy and efficiency criteria. This largely stems from the fact that healthcare data is massive and diverse, encompassing various structured, semi-structured, and unstructured formats like EHRs, medical images, and sensors data (Acharya & Das, 2024; Li et al . , 2021). This, in turn, calls for the need for advanced approach of data pre-processing and feature extraction that would enable the models to learn from the data (Kaur et al. , 2023; Reddy et al. , 2022). Further, many of the devised ML models have drawbacks to include overfitting, scalability challenges, and high computational complexities as associated with large datasets (Lantz, 2019; Oyewola et al. , 2022).

This study aims at using and evaluating the high-end numerics across artificial intelligence for enhancing the healthcare solutions. The proposed work carries out the goal of improving the learnability and convergence of predictive models through the incorporation of gradient descent and Newton's method. For purposes of this research, these numerical methods will reveal how they can solve the issues with big and complicated data in the healthcare setup, which will enhance the efficiency of the conclusions made in health care analytics (Abdel-Basset et al. , 2023; Adam & Mukhtar, 2024).

1. LITERATURE REVIEW

2.1 Predictive Analytics in Healthcare: Current State and Challenges

Future healthcare forecasting was revolutionized with the incorporation of ML and big data. These enhancements have helped the healthcare facilities to analyse big data in terms of predicting the possible outcomes of patients and determining the efficiency of treatment and even organization work (Abdelaziz et al. , 2018; Jayatilake & Ganegoda, 2021). The main areas of focus of big data analytics in the field of healthcare include the determination of patients' risk factors for readmissions, early detection of diseases and efficient allocation of healthcare resources (Fei et al. , 2021; Golas et al. , 2018).

However, several limitations are found in the implementation of predictive analytic in the healthcare industry. A major factor is the sheer amount of healthcare data and possibly, it is also diverse as it can consist of EHRs, image data, sensor data and a lot more into it (Acharya and Das, 2024, Li et al. , 2021). Due to this, there is need for enhanced data preprocessing and feature selection methods so that the models can be able to learn from the data provided (Kaur et al. , 2023; Reddy et al. , 2022). Furthermore, healthcare data particularly possess noise and missing values among other inconvenient factors, which would impact the accuracy of the predictive models (Baghdadi et al. , 2023).

2.2 Numerical Methods: Key Techniques and Their Applications in Machine Learning

Numerical methods are the methods which involve the algorithms of calculations to solve quantitative problems. In the field of machine learning, these techniques are used to find the optimal values of model coefficients, for processing big datasets, and enhancing the computational capability. Numerical methods are such techniques as gradient descent, Newton's method, and optimizers (Abdel-Basset et al. , 2023; Rackauckas et al. , 2020).

Gradient descent is applied in training fully connected neural networks as well as other machine learning algorithms. It keeps on fine-tuning the values of model's parameters in a way that brings the values of the loss function closer to zero thereby enhancing the accuracy of the model itself (Lantz, 2019). The computational advantage of Newton's method is faster convergence because it involves second-order derivatives; it is best suited for fine-tuning of the parameters of the given model where the scenario is complex (Oyewola et al. , 2022). Other forms of optimization are also used to improve feature selection and model training; this include metaheuristic algorithms Kaur et al. , 2023; Ahmed et al. , 2022.

2.3 Integration of Numerical Methods and AI: Previous Studies and Their Outcomes

المجلة العراقية للبحوث الإنسانية والإجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN 2790-1254

Numerical methods have been discussed to have a good synergy when applied in combination with AI and machine learning in healthcare analytics. Scholarly researches have shown that these techniques can help to increase the predictive models precision and speed substantially. For instance, Abdelaziz et al. (2018) used a machine learning model in a cloud computing context and showed a possibility of an improving healthcare services by raising the capability of data processing. In the same regard, Adam and Mukhtar (2024) have also pointed that Ratner et al have observed the amalgamation of AI, machine learning, as well as deep learning in enhancing advanced brain as well as heart care.

In another study, Baghdadi et al. (2023) similarly employed various sophisticated, contemporary machine learning methods for the identification of CVD and showed the viable nature of these methods in the healthcare provisions where they are needed most. Furthermore, Ahmed et al. (2022) combine an optimization and machine learning method to tackle the problem of predicting the emergency patient admission status with remarkable accuracy rates and organizational effectiveness.

Therefore, in healthcare Shahzad et al. (2021) developed a novel optimized predictive framework based on deep learning for healthcare applications, which proves that numerical methods integrated with machine learning are worthy. Further, Pan et al. , (2022) conducted a study to investigate the effect of categorical and numerical attributes in ensemble machine learning architecture to predict the heart disease, thus emphasizing the significance of feature engineering in predictive analytics.

In summary, the combination of numerical methods and machine learning has been witnessed as a strong strategy in the analysis in healthcare with focus on predictive analysis. These techniques are enriching the model performance as well as bearing impacts on the scientific findings and application which in turn helps in providing better patients' care and healthcare results (Abdel-Basset et al. , 2023; Kumar, 2018; Razzak et al. , 2020).

2. METHODOLOGY

3.1 Data Collection

The datasets used in this research were sourced from the publicly available data repositories and Institutional electronic health records (EHRs). Three of these are patient data that encompass; age, gender, ethnicity, past medical history, diagnosis data, laboratory data results, and treatment data. Of records pertaining to different domains, specific ones were selected for diverse predictive ideas like rate of readmission of the patient, identification of cardiovascular diseases at an early stage

المجلة العراقية للبحوث الإنسانية والإجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN 2790-1254

and more importantly, the admission of patients into emergency (Abdelaziz et al. , 2018; Baghdadi et al. , 2023).

3.2 Preprocessing

Data preprocessing is actually a significant part of the data preparation process so as to guarantee the standard of the input data. This process involved several steps:

3.2.1 Data Cleaning

Handling of missing values involved replacing the missing values with either the mean/mode value for the respective variable or through predictive analytics. Censored and conspicuous values were dealt with manageably in a way that they would not influence the findings (Li et al. , 2021).

3.2.2 Normalization and Standardization

In this study, categorical variables were dummified while continuous variables were normalized or standardized due to variations in the differing features to enable proper training of machine learning algorithms (Kaur, et al. , 2023).

3.2.3 Feature Engineering

Some or all of the features were selected based on prior knowledge of the specific domain and or statistical considerations. Other methods such as the Principal Component Analysis (PCA), Recursive Feature Elimination (RFE) were used to achieve dimensionality reduction and boost the model’s performance (Pan et al. , 2022).

3.3 Numerical Methods Applied

To improve the machine learning models' performance, various numerical methods were employed:

3.3.1 Gradient Descent

This strategy was applied during the training of the model when we wanted to reduce the sum of squared errors. Other versions including Stochastic Gradient Descent (SGD) and Mini-Batch Gradient Descent were used to improve on convergence rate and steadiness (Lantz, 2019).

3.3.2 Newton's Method

Print ISSN 2710-0952 Electronic ISSN 2790-1254

In the models where precision was needed Newton's Method was used. The second-order derivatives which are applied in this technique assisted in increasing the speed of convergence and identifying better characteristics of the model parameters (Oyewola et al. , 2022).

3.3.3 Metaheuristic Algorithms

Feature selection methods applied include Genetic Algorithms and Particle Swarm Optimization while the hyperparameters adjustment involved the same methods. These algorithms proved to be useful in mastering the different search spaces and determine the best solutions for the problems (Kaur et al. , 2023; Ahmed et al. , 2022).

3.4 Machine Learning Models

Various machine learning models were selected based on their suitability for different predictive tasks:

3.4.1 Neural Networks

Real time tasks and other deep learning techniques used included CNNs and RNNs on complex pattern analysis, and temporal data respectively by (Adam and Mukhtar in 2024).

3.4.2 Decision Trees and Random Forests

These models were employed due to their ability to process categorical and numerical data with high levels of interpretability and also, high levels of resistance (Lantz, 2019).

3.4.3 Support Vector Machines (SVM)

SVMs were selected intentionally to perform in high-dimensional spaces and benefits from linear relationships (Shahzad et al. , 2021).

3.5 Integration Process

The integration of numerical methods into machine learning models involved several steps:

3.5.1. Model Initialization

The above model was set with prior parameters to commence with the chosen model.

3.5.2. Optimization

Other techniques that have been used include; gradient descent and Newtons method to optimize the model iteratively.

3.5.3. Feature Selection

In this research, metaheuristic algorithms were employed in feature selection to establish the optimal features to process in addition to mitigating the high dimensionality problem that may affect the model.

3.5.4. Training and Validation

Thus the optimized model was built using the preprocessed data and tested using cross validation to check generality of the model(Abdelaziz et al. , 2018).

3.6 Evaluation Metrics

The performance of the machine learning models was evaluated using several metrics:

3.6.1. Accuracy

The correctly predicted instances divided by all the instances.

3.6.2. Precision

The proportion of true positive predictions out of the entire positive predictions made with the attribute to be predicted.

3.6.3. Recall

The proportion of correctly predicted cases of the positive class over the total cases of the positive class.

3.6.4. F1 Score

The F measurements which are the harmonic means of both the precision and the recall which give a balance of the two values.

3.6.5. Area Under the Receiver Operating Characteristic Curve (AUC-ROC)

It measures the extent of capacity where in the model is able to separate the various classes (Sheng et al. , 2021).

Metrics of the evaluation let the models serve their purpose in a balance between precision and recall, and the accuracy of the models was overall good. These metrics were very vital in the evaluation of the numerical methods' effectiveness in

Print ISSN 2710-0952 Electronic ISSN 2790-1254

enhancing the predictive models for the healthcare sector (Jayatilake & Ganegoda, 2021).

4. PRACTICAL ASPECT

4.1 Implementation

Numerical methods are integrated into the machine learning models for healthcare predictive analytics in the following processes. This segment gives practical details on how to use these techniques in an ideal manner.

4.1.1. Data Preparation

- **Data Collection**: Obtain appropriate datasets related to healthcare from areas that may include their own health facilities EHRs and other online databases.

- **Data Preprocessing**: Impute the missing values, check for outliers, and scale the features for better training on the input data. This phase assure the data is prepared in an appropriate form for model training It is necessary to mention that this step is crucial to the next phase, the model training.

4.1.2. Feature Engineering

- **Feature Extraction**: Learn the crucial features of the raw data to be used in the analysis. It may comprise of details such as age, gender, family history, relevant diseases, pathology tests results, and the specifics of treatment.

- **Feature Selection**: There is also the selection of important variables using metaheuristic methods like Genetic Algorithms and Particle Swarm Optimization. This is done to decrease the size of the features space and improve the models' predictive abilities.

4.1.3. Model Selection

- **Neural Networks**: For tasks that require pattern analysis, employ CNN while tasks that have temporal data features should be solved using RNNs.

- **Decision Trees and Random Forests**: These models should be applied for their interpretability as well as when there is both categorical and numeric data involved.

- **Support Vector Machines (SVMs)**: Use SVMs on problems for which you have to live in a high-dimensional space and for non-linear models.

4.1.4. Optimization Techniques

- **Gradient Descent**: Perform gradient descent and the modifications of it such as Stochastic Gradient Descent, Mini-Batch Gradient Descent for the learning of the parameters of the model.
- **Newton's Method**: Use Newton's method when solving models that are more precise and converging faster as compared to the others.
- **Metaheuristic Algorithms**: Use Genetic Algorithms and Particle Swarm Optimization methodologies for hyperparameter optimization, and selecting the features.

4.1.5. Model Training and Validation

- **Training**: Build the selected models on the preprocessed dataset of the chosen optimization techniques for the selected models.
- **Cross-Validation**: Cross validation should be used in order to check the versatility of the model.

4.1.6. Evaluation

- **Performance Metrics**: Evaluate the models based on the performance indices including accuracy, precision, recall, f-measure, and AUC-ROC.

4.2 Case Study: Predicting Patient Readmission Rates

In order to demonstrate an example of the usage of these methodologies, a case was done with focus on patient readmission rates. The following steps outline the process:

4.2.1. Data Collection and Preprocessing

- Electronic health records from an institutional source were used to obtain patient data and this included demographic details of the patients, their clinical history and treatment details.
- First, the data was cleansed and standardized; also, only relevant features were selected.

4.2.2. Feature Engineering

- Metaheuristic algorithms were used to define the most significant features influencing the readmission rates.

4.2.3. Model Selection and Optimization

- A Random Forest model was chosen due to its interpretability and, at the same time, high robustness.
- Gradient Descent in addition to Genetic Algorithms were deployed for the purpose of optimizing model parameters as well as selecting the most appropriate features.

4.2.4. Training and Validation

- Finally, the model was learned on the preprocessed data and validated employing the cross validation methodology.

4.2.5. Evaluation

- The performance metrics used to quantify the findings of the model are accuracy, precision, recall, F1 score, and AUC-ROC. In analyzing the findings of the study, it was observed that the optimized model performed well in forecasting the readmission rates of the patients compared to other baseline models.

4.3 Results

The use of numerical methods posed a boost to the performance of machine learning models in the healthcare predictive analytics. Key findings include:

4.3.1. Enhanced Accuracy

The models optimized for target classes demonstrated better predictive capability of patient’s status.

4.3.2. Improved Efficiency

The use of integrated numerical methods ensured that the time spent in training the model as well as optimizing the models was minimized.

4.3.3. Better Generalizability

Specifically, the models showed better being able to generalize from the single dataset used for creating the model, evident by their performance on cross-validation datasets.

4.4 Discussion

This is evident from the results obtained in improving the performance of the areas of applications for healthcare through the use of numerical methods in the field of machine learning. As the results show it, these methods may be useful for improving the accuracy, thus enhancing patient outcomes and increasing the general effectiveness of functioning in a healthcare organization. However, the following negative aspects should also be taken into account: in order to calculate the solutions, a large amount of computational operations may be required and the implementation of high-level methods of numerical analysis may also be problematic.

4.5 Limitations

While the integration of numerical methods showed significant improvements, some limitations need to be addressed:

4.5.1. Computational Complexity

Higher numerical techniques may have many computational demands on the healthcare facilities, and this may be regarded as a limitation.

4.5.2. Data Quality

Depending on the availability and purity of the data, Biohealthcare is highly vulnerable to the model's accuracy. It means that when the data are not complete or not representative of the total population, forecasts are equally not going to perfect.

4.5.3. Generalizability

While cross-validation is useful in evaluating generalization applicability can be somewhat different depending on a particular healthcare setting and the patients, included.

4.6 Future Work

The limitations of the study can be further addressed in future research through identifying better algorithms that can enhance the results of data analysis with the help of improved computational tools. Also, moving beyond the examples of healthcare personnel workload, the use of numerical methods in case studies of other healthcare applications can add more confidence to the usability of solutions and predictive analytics.

Thus, in terms of practical and theoretical work, the quantitative methods for analyzing how machine learning in health care can be improved, are useful for the continuous advancement of predictive analytics which benefits to enhance patients' results and experiences as well as improvements in the related services.

5. CONCLUSION

العدد 15 تشرين2 2024
No.15 Nov 2024
المجلة العراقية للبحوث الإنسانية والإجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN 2790-1254

The subject of this research focused on the improvement of machine learning models for predictive analytics in healthcare through enhanced numerical method incorporation. The first goal was to improve predictive models to existing large and complicated healthcare database issues and opportunities.

5.1 Summary of Findings

5.1.1. Enhanced Model Performance

Numerical methods for the optimization of parameters of machine learning models like gradient descent, Newton's method, and metaheuristic algorithms increased the efficiency of the models. Specificity enhancements could also be observed in figures like accuracy, precision, recall, F1 score, and AUC-ROC.

5.1.2. Efficient Optimization

The incorporation of optimization approaches facilitated the speed of the model development especially in terms of training and hyper-parameter selection.

5.1.3. Better Generalizability

Cross-validation results also shown that the models optimized with numerical methods were more generalizable, perhaps suggesting that these approaches could be applied to other health care datasets easily.

5.1.4. Practical Implementation

The usefulness of these methods can be explained by a case study on predicting patient's readmission rates for patients with heart failure. The baseline models as well as the optimized model confirmed that the study had practical real life application.

5.2 Impact on Healthcare

The findings of this study have significant implications for healthcare:

5.2.1. Improved Patient Outcomes

Thus, highly accurate models can be of tremendous help for providers who need to make sound decisions when it comes to the patient's health and care.

5.2.2. Operational Efficiency

Optimizing the model means that the time and the money being spent on the analysis can be less thereby increasing the efficiency of the healthcare institutions.

Print ISSN 2710-0952 Electronic ISSN 2790-1254

5.2.3. Scalability

The techniques that have been discussed above can be applied on different predictive problems that are prevalent in healthcare; therefore the proposed approach is very flexible.

5.3 Limitations

Despite the promising results, the study faced several limitations:

5.3.1. Computational Demands

Some of these advanced numerical methods demand a lot of computation which most centres of healthcare delivery may not afford.

5.3.2. Data Quality and Availability

Its important in selection of adequate model and quality and completeness of health care data for prediction. When data is not complete or when only a part of relevant data is used, the predictions, which are derived out of it, are likely to be going wrong.

5.3.3. Complex Implementation

Subsequently, fine-tuning as well as the identification of suitable numerical methods can be time-consuming and involve the need for excessive technical knowledge.

5.4 Future Work

The following are some of the methods that can be suggested for improvement as the future study spots on a more efficient and accessible algorithm: Potential directions for further study include:

5.4.1. Algorithmic Advancements

Improving existing numerical methods that can be solved with less computational power and creating new ones.

5.4.2. Broader Applications

Expanding the study to the other predictive tasks to more healthcare domains to determine the transferability of the techniques.

5.4.3. Data Enhancement

Studying on the concepts of enhancing the data quality and data completeness for better model prediction.

المجلة العراقية للبحوث الإنسانية والإجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN 2790-1254

Thus, the incorporation of higher numerical calculations into machine learning-based algorithms can be considered as the successful avenue for enhancing predictive visualization in medicine. This work shows the possibility of enhancement in accuracy, speed and interpretability of the constructed models that can lead to enhancement of the global functioning of healthcare systems and consequently improvement of the patient's health. The ongoing work in this area will only provide for the improvement of the potentials and effectiveness of predictive analytics in the application area of interest which is healthcare.

REFERENCES

Abdelaziz, A., Elhoseny, M., Salama, A. S., & Riad, A. M. (2018). A machine learning model for improving healthcare services on cloud computing environment. *Measurement*, *119*, 117-128.

Abdel-Basset, M., Chakrabortty, R. K., & Mohamed, R. (2023). *Application of Advanced Optimization Techniques for Healthcare Analytics*. CRC Press.

Acharya, S., & Das, S. (2024). Statistical Modeling for Predictive Healthcare Analytics. In *Revolutionizing Healthcare Treatment With Sensor Technology* (pp. 265-293). IGI Global.

Adam, M. A., & Mukhtar, A. (2024). Smart Health Solutions: The Convergence of AI, Machine Learning, and Deep Learning for Brain and Heart Care. *Revista Espanola de Documentacion Cientifica*, *18*(02), 238-268.

Ahmed, A., Ashour, O., Ali, H., & Firouz, M. (2022). An integrated optimization and machine learning approach to predict the admission status of emergency patients. *Expert Systems with Applications*, *202*, 117314.

Baghdadi, N. A., Farghaly Abdelaliem, S. M., Malki, A., Gad, I., Ewis, A., & Atlam, E. (2023). Advanced machine learning techniques for cardiovascular disease early detection and diagnosis. *Journal of Big Data*, *10*(1), 144.

Fei, Z., Ryeznik, Y., Sverdlov, O., Tan, C. W., & Wong, W. K. (2021). An overview of healthcare data analytics with applications to the COVID-19 pandemic. *IEEE Transactions on Big Data*, *8*(6), 1463-1480.

Golas, S. B., Shibahara, T., Agboola, S., Otaki, H., Sato, J., Nakae, T., ... & Jethwani, K. (2018). A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. *BMC medical informatics and decision making*, *18*, 1-17.

Jayatilake, S. M. D. A. C., & Ganegoda, G. U. (2021). Involvement of machine learning tools in healthcare decision making. *Journal of healthcare engineering*, *2021*(1), 6679512.

Kaur, S., Kumar, Y., Koul, A., & Kumar Kamboj, S. (2023). A systematic review on metaheuristic optimization techniques for feature selections in disease diagnosis: open issues and challenges. *Archives of Computational Methods in Engineering*, *30*(3), 1863-1895.

Kumar, V. V. (2018). *Healthcare analytics made simple: techniques in healthcare computing using machine learning and Python*. Packt Publishing Ltd.

Lantz, B. (2019). *Machine learning with R: expert techniques for predictive modeling*. Packt publishing ltd.

Li, W., Chai, Y., Khan, F., Jan, S. R. U., Verma, S., Menon, V. G., ... & Li, X. (2021). A comprehensive survey on machine learning-based big data analytics for IoT-enabled smart healthcare system. *Mobile networks and applications*, *26*, 234-252.

Ngiam, K. Y., & Khor, W. (2019). Big data and machine learning algorithms for health-care delivery. *The Lancet Oncology*, *20*(5), e262-e273.

Ogunpola, A., Saeed, F., Basurra, S., Albarrak, A. M., & Qasem, S. N. (2024). Machine learning-based predictive models for detection of cardiovascular diseases. *Diagnostics*, *14*(2), 144.

Oyewola, D. O., Dada, E. G., Omotehinwa, T. O., Emebo, O., & Oluwagbemi, O. O. (2022). Application of deep learning techniques and Bayesian optimization with tree parzen Estimator in the classification of supply chain pricing datasets of health medications. *Applied Sciences*, *12*(19), 10166.

Pan, C., Poddar, A., Mukherjee, R., & Ray, A. K. (2022). Impact of categorical and numerical features in ensemble machine learning frameworks for heart disease prediction. *Biomedical Signal Processing and Control*, *76*, 103666.

Rackauckas, C., Ma, Y., Noack, A., Dixit, V., Mogensen, P. K., Byrne, S., ... & Ivaturi, V. (2020). Accelerated predictive healthcare analytics with pumas, a high performance pharmaceutical modeling and simulation platform. *BioRxiv*, 2020-11.

Ratner, B. (2017). *Statistical and machine-learning data mining:: Techniques for better predictive modeling and analysis of big data*. Chapman and Hall/CRC.

Razzak, M. I., Imran, M., & Xu, G. (2020). Big data analytics for preventive medicine. *Neural Computing and Applications*, *32*(9), 4417-4451.

Reddy, D. K. K., Behera, H. S., & Ding, W. (2022). Advance Machine Learning and Nature-Inspired Optimization in Heart Failure Clinical Records Dataset. In Nature-Inspired Optimization Methodologies in Biomedical and Healthcare (pp. 221-246). Cham: Springer International Publishing.

Santamato, V., Tricase, C., Faccilongo, N., Marengo, A., & Pange, J. (2024). Healthcare performance analytics based on the novel PDA methodology for assessment of efficiency and perceived quality outcomes: A machine learning approach. *Expert Systems with Applications*, *252*, 124020.

Shahzad, Y., Javed, H., Farman, H., Ahmad, J., Jan, B., & Nassani, A. A. (2021). Optimized Predictive Framework for Healthcare Through Deep Learning. *Computers, Materials & Continua*, *67*(2).

Sheng, J. Q., Hu, P. J. H., Liu, X., Huang, T. S., & Chen, Y. H. (2021). Predictive analytics for care and management of patients with acute diseases: deep learning–based method to predict crucial complication phenotypes. *Journal of medical Internet research*, *23*(2), e18372.

Tawhid, A., Teotia, T., & Elmiligi, H. (2021). Machine learning for optimizing healthcare resources. In *Machine Learning, Big Data, and IoT for Medical Informatics* (pp. 215-239). Academic Press.

Thethi, S. K. (2024). Machine learning models for cost-effective healthcare delivery systems: A global perspective. *Digital Transformation in Healthcare 5.0: Volume 1: IoT, AI and Digital Twin*, 199.

Waring, J., Lindvall, C., & Umeton, R. (2020). Automated machine learning: Review of the state-of-the-art and opportunities for healthcare. *Artificial intelligence in medicine*, *104*, 101822.

Xiao, C. (2016). *Optimization and Machine Learning Methods for Medical and Healthcare Applications* (Doctoral dissertation).

Zhai, K., Yousef, M. S., Mohammed, S., Al-Dewik, N. I., & Qoronfleh, M. W. (2023). Optimizing clinical workflow using precision medicine and advanced data analytics. *Processes*, *11*(3), 939.