Systems Assessment and Engineering Management

Journal Homepage: sciencesforce.com/saem

SCIENCES FORCE
International Academic Publisher

Syst. Assess. Eng. Manage. Vol. 2 (2024) 23–32

Paper Type: Original Article

A Fair Approach to Heart Disease Prediction: Leveraging Machine Learning Model

Rizwan Karim [1,*] and Muhammad Imran Asjad [1]

[1] Department of Mathematics, University of Management and Technology, Lahore, Pakistan.
Emails: rizwankarim3@gmail.com; imran.asjad@umt.edu.pk.

Received: 27 Aug 2024 **Revised:** 06 Nov 2024 **Accepted:** 30 Nov 2024 **Published:** 01 Dec 2024

Abstract

This study focuses on the tasks of diagnosing and predicting diseases, which are crucial, for accurately classifying and treating them by cardiologists. By utilizing the increasing use of machine learning in the field in pattern recognition from data this research introduces a specialized model that aims to predict cardiovascular diseases. The main objectives of this model are to reduce misdiagnosis rates and minimize fatalities. To achieve these goals the proposed approach combines Logistic Regression with a fairness component. The model is trained using a real world dataset consisting of 70,000 instances obtained from Kaggle. The dataset is split into 70% for training and 30% for testing purposes to evaluate accuracy and fairness metrics at values of Logistic Regression. Through reweighing techniques applied to the model improvements, in both accuracy and fairness are observed. In conclusion this research suggests that machine learning models that prioritize fairness demonstrate performance by achieving an accuracy rate of 72% with a fairness value of 0.009.

Keywords: Fairness; Machine Learning; Heart Disease; Classification; Model Evaluation.

1 | Introduction

Cardiovascular disease (CVD) is a significant global health issue since it is the main cause of illness and death worldwide. It is responsible for more than 70% of all fatalities. CVD accounts for approximately 43% of all mortality, according to the 2017 Global Burden of Disease study [1, 2]. In high-income nations, a variety of common risk factors contribute to the prevalence of heart disease. Poor dietary choices, tobacco use, excessive sugar consumption, obesity, and excess body fat are among these risk factors [3]. Such risk factors are a key contributor to the high prevalence of cardiovascular disease in these nations. Importantly, the burden of CVD is not exclusive to high-income countries. Chronic illnesses, particularly cardiovascular disorders, are becoming more common in low- and middle-income nations. This implies that the problem is not limited to wealthy countries and emphasizes the global aspect of the problem. To provide some further context, it was predicted that the staggering USD 3.7 trillion worldwide economic burden of cardiovascular illnesses will be attained between 2010 and 2015 [4, 5]. This massive financial load has far-reaching repercussions for healthcare systems, economies, and civilizations globally.

In recent years, there has been a substantial increase in research utilizing machine learning (ML) to improve healthcare solutions [6-9]. However, as interest in ML applications develops, concerns have developed about

Corresponding Author: rizwankarim3@gmail.com
https://doi.org/10.61356/j.saem.2024.2438

the potential for these technologies to increase current health inequities. The term "fair ML" has emerged as a crucial issue, with the primary objective of creating strategies to prevent ML models from disproportionately hurting existing marginalized and excluded people. One of the most difficult challenges in the field of fair ML is defining fairness and inequity in the context of machine learning. In ML, the quest for fairness frequently focuses on minimizing differences in model performance while taking into account subsets of demographic variables [10].

Electrocardiograms and CT scans are examples of expensive and ineffective diagnostic techniques that have led to the deaths of 17 million people [11]. Furthermore, employees suffering from cardiovascular disorders account for a major amount of a company's yearly medical expenditures, accounting for around $25-30\%$ of annual medical expenses. It is critical to emphasize early identification as a method of reducing the physical and financial toll that heart disease puts on individuals and organizations. As per the World Health Organization (WHO), it is projected that the global mortality attributed to cardiovascular diseases will increase to 23.6 million by the year 2030, with heart disease and stroke identified as the main contributing factors [12].

Fairness in machine learning is crucial in our effort to help save lives and minimize the societal cost of healthcare [13]. We can make prediction tools for heart disease risk assessment more equal and reasonable by including fairness factors in data mining and machine learning algorithms. This not only aids in early detection but also helps to eliminate biases that may disproportionately harm particular demographic groups, eventually encouraging better healthcare results for everybody.

Cardiovascular disease (CVD), often known as heart disease, is a major global health concern, with more than 70% of all worldwide deaths [14]. Circulatory illness accounts for about 43% of all deaths, according to the 2017 Global Burden of Illness Study [15]. These figures highlight the enormous impact of heart disease on public health. This disease is caused by a variety of risk factors, many of which are prominent in high-income nations. Unhealthy diets, tobacco use, excessive sugar consumption, and being overweight or obese are major causes. It is important to emphasize, however, that low- and middle-income nations are not immune to the increased frequency of chronic illnesses, including CVDs [16].

Additionally, key diagnostic procedures like EKGs (electrocardiograms) and CT scans, which are crucial for diagnosing coronary heart disease, frequently create financial and practical difficulties, especially in low- and middle-income countries [17]. As a result, early detection of heart conditions is crucial for reducing the physical and financial strain on individuals and organizations.

Machine learning is crucial in the field of healthcare. It enables us to identify, find, and predict a variety of medical disorders. In recent years, there has been increasing interest in utilizing data mining and machine learning technology to forecast the likelihood of developing specific diseases [18-21]. Although data mining approaches have been investigated in the past to predict illnesses, certain studies have had difficulty properly predicting the danger of the disease progressing. The fundamental goal of this work, in this context, is to improve the precision of forecasting the possible incidence of heart disease in individuals, following the principles of fairness in machine learning.

While machine learning algorithms have achieved significant advances in prediction accuracy, questions about interpretability and fairness remain [22]. These qualities are critical in light of their upcoming deployment in translational research and practical application. Historically, machine learning research in healthcare has mostly focused on evaluating the model's overall performance. This evaluation was primarily concerned with determining how well the model predicted preset outcomes inside a test dataset. However, there is a changing paradigm in recent times, characterized by growing anxiety about the success of these models, particularly concerning under-represented and marginalized demographic groups. These populations have raised increasing amounts of worry since they were often under-represented in the training data used to create the models. These issues highlight the vital relevance of addressing any inequities and biases that may occur when deploying population-level predictive models, a significant problem for medical professionals involved in the implementation of such models. In this context, concerns about the idea of "fairness" in machine learning relate to the issue of algorithmic bias in machine learning approaches [23]. Algorithmic bias is the tendency of machine learning models to consistently predict certain outcomes with a higher likelihood for one group versus another, especially when these groupings are based on attributes that are considered sensitive and should ideally bear no significant correlation with the expected outcomes. These characteristics might include, among other things, gender, age, weight, cholesterol levels, and glucose levels. When we deploy a model that has strong predictive performance in the general population but is intrinsically biased against marginalized or underprivileged groups, we have reason to be concerned. This prejudice may have negative repercussions for patients in these underserved subcohorts. In summary, it poses the unsettling potential of perpetuating inequities in healthcare and increasing current inequalities, which is a topic of major relevance and ethical issue in the field of machine learning applications in healthcare.

Closely related work by other researchers focusing on the use of machine learning models to diagnose and predict cardiovascular diseases. Many studies [24-27] in the medical and machine learning fields have explored similar objectives, aiming to develop accurate predictive models to aid in diagnosis and treatment, however, there are key difference in the mentioned studies is to incorporate a component of justice into the machine learning model. This highlighted lack of bias in predictive models is less common in traditional medical literature, where the primary focus tends to be that it would provide greater accuracy than dealing with potential biases or inaccuracies in model forecasts would do so or not use objectivity-enhancing methods such as additional weighting methods. This study is therefore differentiated by its twin objectives of accuracy and unbiasedness, which strive to lower the rate of misdiagnosis and mortality, while also prioritizing fairness and predictions.

As of now, there is a notable lack of comprehensive assessments explicitly dedicated to analyzing the idea of fairness in machine learning models used to forecast heart disease. To close this huge gap, the current work conducts a retrospective cohort analysis with the primary goal of examining the fairness of a cutting-edge machine learning model known for its remarkable performance in predicting acute cardiac events. The basic premise underlying this work is that, even within models that display outstanding overall competency in terms of classification accuracy, noticeable inequalities across specific subpopulations may remain. These subpopulations can be characterized by a specific trait or characteristic, such as a patient's degree of physical activity (whether active or not). The objective of this hypothesis is to throw light on the possibility of biases or imbalances in model predictions that may have previously gone unreported due to a solitary concentration on general forecast accuracy. We hope to add to the expanding body of knowledge on the equitable and ethical implications of machine learning applications in the domain of heart disease prediction by researching and comprehending these inequalities.

2 | Data Set

The process of creating the training and test datasets for our cardiovascular disease risk prediction model with fairness included the use of a CVD-specific dataset. This CVD dataset is an open-source resource gathered from Kaggle, with a total of 70,000 participants. This dataset's balance is an important feature since it includes about an equal number of people who are healthy and those who have been diagnosed with cardiovascular disorders. Eleven different properties in this dataset serve as critical variables for our prediction model. These characteristics are divided into three categories: demographic, examination, and social history. Age (in years) and gender are two demographic characteristics. Weight (measured in kilograms) and height (measured in square meters) are examination-related characteristics that allow us to calculate the Body Mass Index (BMI) using the formula $\mathrm{BMI} = \frac{\text{weight}}{(\text{ height })^2}$. This dataset also includes cholesterol and glucose values, which are classified as normal, above normal, or far above normal. Physical activity, drinking habits (with values 'yes' or 'no'), and smoking habits (also with values 'yes' or 'no') are among the social history-related characteristics. Furthermore, the dataset contains systolic and diastolic blood pressure measurements in millimeters of mercury (mmHg), which were obtained during the subjects' medical examinations.

3 | Methodology

The major goal of this project is to create an advanced system for forecasting the likelihood of heart disease utilizing cutting-edge computerized heart disease prediction algorithms, with a special emphasis on fairness. The suggested approach has the potential to provide major advantages to both medical professionals and patients. To achieve this ultimate aim, we employed different machine learning classification models to make predictions on heart disease data while ensuring fairness in our approach. This study report contains the outcomes of this extensive research project. We thoroughly prepared the dataset to ensure the robustness of our methods. This includes data cleansing to remove any inconsistencies or errors, as well as the removal of any unnecessary or irrelevant information. Certain attributes have been standardized to better improve the data. For instance, we turned age into years and weight into kilograms. These standardizations improve the data's homogeneity and precision, which contributes to the model's overall quality. Most importantly, our research emphasizes fairness in the prediction model. To do this, we have labeled some groups as privileged and unprivileged, notably in terms of a protected quality - the gender designation in Connection with heart health. This is done to guarantee that our model is neutral and equitable in its predictions, taking into consideration any potential discrepancies that may occur across various demographic groupings. After all of this, we proceed to train our model on the processed dataset. We predict that by using this new process, we will get not just more accurate findings, but also greater overall model performance. This is objectively proven in Figure 1, emphasizing the efficacy and dependability of our technique.

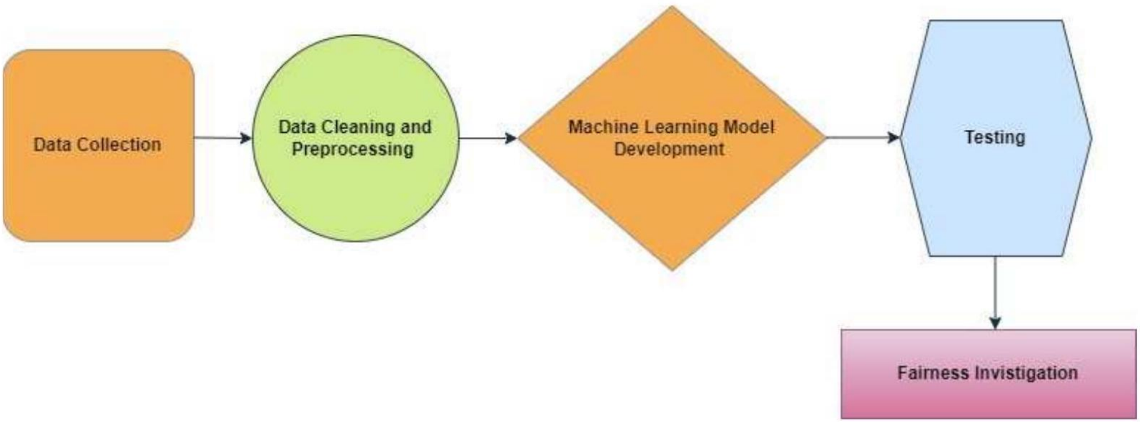


Figure 1. Workflow.

3.1 | Machine Learning Model

In our study, we used a Logistic regression classifier to predict cardiology outcomes. We considered gender as a factor, for all patients. The overall accuracy we achieved was 0.7171. There was a small difference of 0.005 in equal opportunity. Figures 2 and 3 visually and Table 2 represent the ranking of accuracies in relation, to opportunity differences at different points. After adjusting the model with a hyperparameter value of 1 we noticed an improvement, in accuracy reaching 0.7211. Moreover, the fairness metric also showed enhancement reaching 0.0019. However, when we changed the hyperparameter to its value of 0.00001 we observed a decrease in model accuracy, to 0.64 and a fairness metric of 0.1504 as shown in Table 3. By using reweighting techniques, in conjunction with regression, we were able to restore the models' accuracy to 0.7183. Additionally, we observed an equal opportunity difference of -0.008. Figures 4 and 5 and Table 4 give a summary of the accuracy rankings. Equal opportunity differences, after applying the reweighting technique with the logistic regression model. After conducting tests and modifying the hyperreweighing value to 1 we achieved an accuracy rate of 0.7217 and a fairness metric of 0.008. On the other hand, when we reduced the hyperreweighing value to 0.00001 both accuracy 0.6436 and fairness 0.0077 decreased as shown in Table 5. These findings highlight how sensitive the models' performance is, to adjustments in hyperparameters and underscore the importance of maintaining a balance, between accuracy and fairness when using modeling.

4 | Fairness

4.1 | Investigating Bias

Our investigation focused on the study of model bias in relation to heart disease characteristics, especially cardio. In this context, we classified groups as privileged or unprivileged based on gender and the relative frequency of cardiac problems. Table 1 gives a detailed description of the protected attribute classes together with the privileged and unprivileged values associated with them. We utilized the Python library "FairLearn" to assess and measure the fairness of our model.

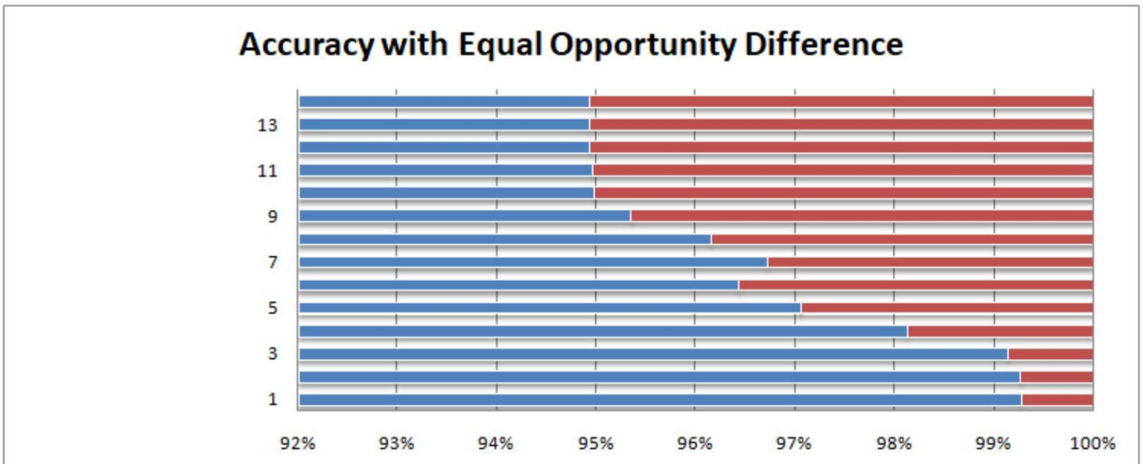
Table 1. A detailed description of the protected attribute classes.

Protected Attribute	Privileged Group	Unprivileged Group
Gender	Male	Female

Our method entailed computing a critical ratio, as shown in Eq. (1), for each unprivileged class and each performance parameter connected with our models. This ratio is significant since it can reflect the degree of fairness in the model's performance [28]. When the ratio approaches one, it indicates that the model performed rather well. To determine the presence of bias in our models across domains, we used the parameter, which has values ranging from 0 to 1 [29]. In our investigation, we used the value 0.7 as a critical threshold to discover bias in the performance of our models, following a well-accepted heuristic known as the "70% rule" [30]. We were able to determine an acceptable range for the model performance ratio, which was between 0.7 and 1.15, by using $\epsilon = 0.7$. This means that if the ratio calculated by Eq. (1) falls within this range (0.7 to 1.15), the model was declared impartial concerning that specific measure and the attribute class associated with it.

$$\forall_{i\varepsilon a,b,c,\dots,z}$$

$$\varepsilon < \frac{metric_i}{metric_{privileged}} < \frac{1}{\varepsilon} \tag{1}$$



	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Accuracy	0.7171	0.7158	0.7084	0.6834	0.6612	0.6521	0.6484	0.6464	0.6461	0.6460	0.6459	0.6459	0.6459	0.6459
Equal opportunity Difference	0.0052	0.0052	0.0061	0.0129	0.0200	0.0240	0.0219	0.0258	0.0315	0.0341	0.0343	0.0345	0.0345	0.0345

Figure 2. Accuracy and Fairness.

Table 2. Accuracy and Fairness.

Accuracy	Equal Opportunity Difference
0.717183673	0.005277534
0.715836735	0.005277534
0.70844898	0.006161339
0.68344898	0.012991104
0.661285714	0.02006145
0.652142857	0.024089961
0.64644898	0.021955587
0.648408163	0.02588629
0.646122449	0.031516424
0.646020408	0.034136513
0.645979592	0.034306022
0.645979592	0.034501114
0.645979592	0.034501114
0.645979592	0.034501114

4.2 | Mitigating Bias

We used the reweighing method to reduce algorithmic bias in our prediction models. During the model's training phase, this strategy included fine-tuning observation weights inside each attribute outcome combination [30-32]. The degree of bias in these newly modified models was then evaluated and contrasted with the bias noted in the initial base models [33]. The reweighing strategy entailed the development of a new model in which observation weights were determined by taking into account the distribution of observations within both unprivileged and privileged categories. This weight adjustment process attempted to re-balance the model's training data to address and correct any prior biases, hence improving the fairness and equity of the model's predictions [34].

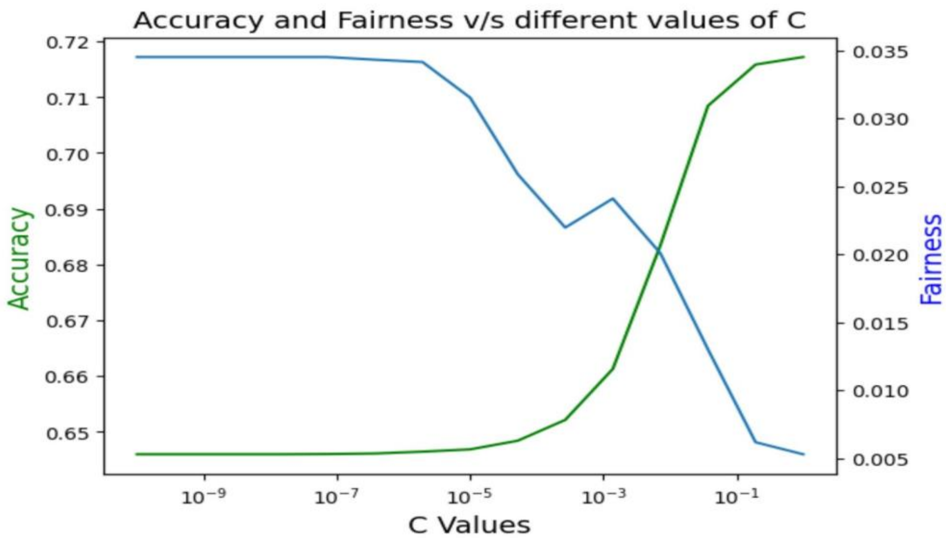


Figure 3. Accuracy and Fairness.

Table 3. Accuracy and fairness with different values of test hyper.

	Test hyper value with 1	Test hyper value with 0.00001
Accuracy	0.721142857142857	0.643809523809523
Fairness	0.00194496537012878	0.0150441223851379

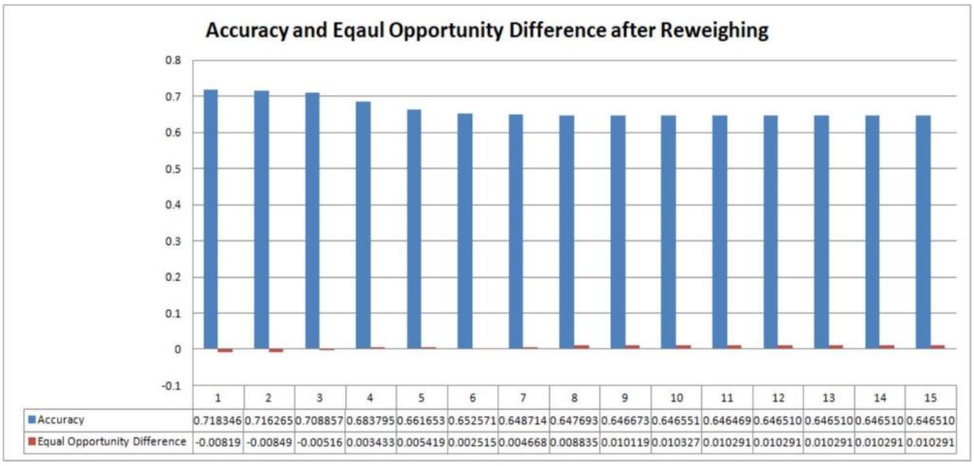


Figure 4. Accuracy and Fairness.

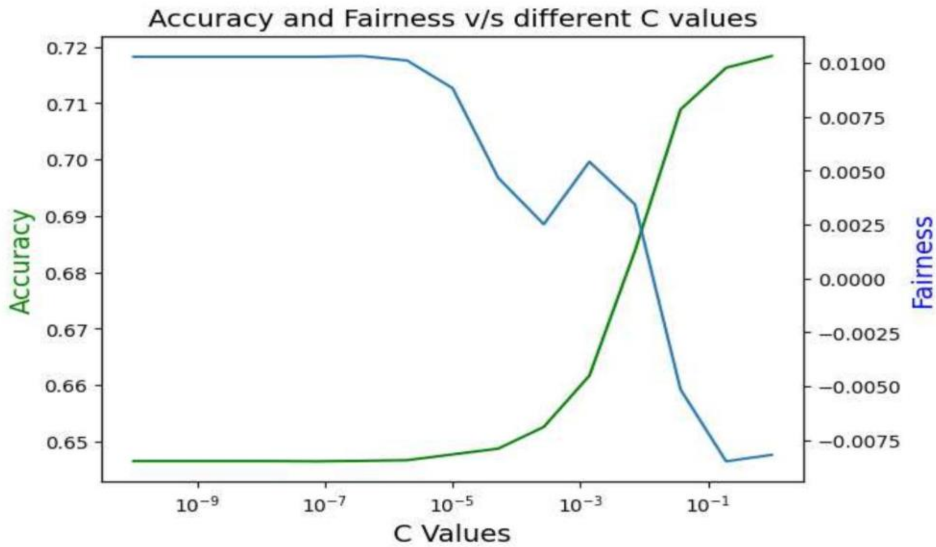


Figure 5. Accuracy and Fairness.

Table 4. Accuracy and Fairness after reweighting.

Accuracy	Equal Opportunity Difference
0.7183469387755103	-0.008192418935764167
0.716265306122449	-0.008494224134953265
0.7088571428571429	-0.005159675782221917
0.6837959183673469	0.003433387591332204
0.6616530612244899	0.005419313078510535
0.6525714285714286	0.002515444635538411
0.6487142857142857	0.0046679897255002215
0.6476938775510204	0.008834988698378222
0.6466734693877552	0.0101193533930096
0.6465510204081631	0.01032761743902828
0.646469387755102	0.010291858677693578
0.6465102040816326	0.010291858677693578
0.6465102040816326	0.010291858677693578
0.6465102040816326	0.010291858677693578
0.6465102040816326	0.010291858677693578

Table 5. Accuracy and Fairness after reweighting.

	Test hyper value of reweighing with 1	**Test hyper value of reweighing with 10^{-4}**
Accuracy	0.7217142857142858	0.6436666666666667
Fairness	-0.008418508646065637	-0.007783519951961804

5 | Conclusion

In summary, this research highlights the importance of incorporating fairness considerations, into machine learning models for predicting diseases. The proposed Logistic Regression model, combined with a fairness component shows promising outcomes in improving both classification accuracy and fairness. The study, conducted on a real-world dataset indicates that adjusting the model at different test hyper values leads to adaptability and better accuracy and fairness metrics. Ultimately these findings support the claim that machine learning models that prioritize fairness as exemplified by the Logistic Regression approach in this study offer a fit by achieving a 72% accuracy alongside a commendable fairness value of 0.009. These results underscore the potential of integrated machine learning models to significantly contribute to diagnosing and predicting diseases thereby advancing patient care, in the field of medicine.

Acknowledgments

All the authors are obliged and thankful to the University of Management and Technology Lahore, Pakistan, for facilitating and supporting the research work.

Funding

This research has no funding source.

Author Contribution

Conceptualization, R.K.; methodology, R.K.; validation, M.A.I.; investigation, M.I.A, R.K; writing original draft preparation, R.K; writing review and editing, M.A.I; visualization, R.K.; supervision, M.I.A.; project administration, M.I.A.; funding acquisition, M.I.A. All authors have read and agreed to the published version of the manuscript.

Data Availability

All data produced or scrutinized in this study has been sourced from the open-access platform Kaggle. Link: https:www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset

Conflicts of Interest

The authors declare that there is no conflict of interest in the research.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

[1] Estes, C., Anstee, Q. M., Arias-Loste, M. T., Bantel, H., Bellentani, S., Caballeria, J., ... & Razavi, H. (2018). Modeling nafld disease burden in china, france, germany, italy, japan, spain, united kingdom, and united states for the period 20162030. Journal of hepatology, 69(4), 896-904.

[2] Drozdz, K., Nabrdalik, K., Kwiendacz, H., Hendel, M., Olejarz, A., Tomasik, A., ... & Lip, G. Y. (2022). Risk factors for cardiovascular disease in patients with metabolic-associated fatty liver disease: a machine learning approach. Cardiovascular Diabetology, 21(1), 240.

[3] Murthy, H. N., & Meenakshi, M. (2014, November). Dimensionality reduction using neuro-genetic approach for early prediction of coronary heart disease. In International conference on circuits, communication, control and computing (pp. 329-332). IEEE.

[4] Mozaffarian, D., Benjamin, E. J., Go, A. S., Arnett, D. K., Blaha, M. J., Cushman, M., ... & Turner, M. B. (2016). Heart disease and stroke statistics2016 update: a report from the American Heart Association. circulation, 133(4), e38-e360.

[5] Maiga, J., & Hungilo, G. G. (2019, October). Comparison of machine learning models in prediction of cardiovascular disease using health record data. In 2019 international conference on informatics, multimedia, cyber and information system (ICIMCIS) (pp. 45-48). IEEE.

[6] Wiens, J., & Shenoy, E. S. (2018). Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. Clinical infectious diseases, 66(1), 149-153.

[7] Javaid, M., Haleem, A., Singh, R. P., Suman, R., & Rab, S. (2022). Significance of machine learning in healthcare: Features, pillars and applications. International Journal of Intelligent Networks, 3, 58-73.

[8] Alanazi, A. (2022). Using machine learning for healthcare challenges and opportunities. Informatics in Medicine Unlocked, 30, 100924.

[9] Jabbar, M. A., Samreen, S., & Aluvalu, R. (2018). The future of health care: Machine learning. International Journal of Engineering & Technology, 7(4.6), 23-25.

[10] Gil, N. L. M. (2022). Minimax Fairness in Machine Learning (Doctoral dissertation, Duke University).

[11] Bhatt, C. M., Patel, P., Ghetia, T., & Mazzeo, P. L. (2023). Effective heart disease prediction using machine learning techniques. Algorithms, 16(2), 88.

[12] Balbay, Y., Gagnon-Arpin, I., Malhan, S., Öksz, M. E., Sutherland, G., Dobrescu, A., ... & Habib, M. (2018). Modeling the burden of cardiovascular disease in Turkey. Anatolian Journal of Cardiology, 20(4), 235.

[13] Davoudi, A., Sajdeya, R., Ison, R., Hagen, J., Rashidi, P., Price, C. C., & Tighe, P. J. (2023). Fairness in the prediction of acute postoperative pain using machine learning models. Frontiers in Digital Health, 4, 970281.

[14] Jagannathan, R., Patel, S. A., Ali, M. K., & Narayan, K. V. (2019). Global updates on cardiovascular disease mortality trends and attribution of traditional risk factors. Current diabetes reports, 19, 1-12.

[15] Amini, M., Zayeri, F., & Salehi, M. (2021). Trend analysis of cardiovascular disease mortality, incidence, and mortality-to-incidence ratio: results from global burden of disease study 2017. BMC Public Health, 21(1), 1-12.

[16] Samb, B., Desai, N., Nishtar, S., Mendis, S., Bekedam, H., Wright, A., ... & Etienne, C. (2010). Prevention and management of chronic disease: a litmus test for health-systems strengthening in low-income and middle-income countries. The Lancet, 376(9754), 1785-1797.

[17] Bonow, R. O., Mann, D. L., Zipes, D. P., & Libby, P. (2011). Braunwald's heart disease e-book: A textbook of cardiovascular medicine. Elsevier Health Sciences.

[18] Chaurasia, D. V., & Pal, S. (2013). Early prediction of heart diseases using data mining techniques. Caribbean journal of Science and Technology, 1, 208-217.

[19] Srinivas, K., Rani, B. K., & Govrdhan, A. (2010). Applications of data mining techniques in healthcare and prediction of heart attacks. International Journal on Computer Science and Engineering (IJCSE), 2(02), 250-255.

[20] Banaee, H., Ahmed, M. U., & Loutfi, A. (2013). Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges. Sensors, 13(12), 17472-17500.

[21] Amin, M. S., Chiam, Y. K., & Varathan, K. D. (2019). Identification of significant features and data mining techniques in predicting heart disease. Telematics and Informatics, 36, 82-93.

[22] Barocas, S., Hardt, M., & Narayanan, A. (2017). Fairness in machine learning. Nips tutorial, 1, 2017.

[23] Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM computing surveys (CSUR), 54(6), 1-35.

[24] Ma, C. Y., Luo, Y. M., Zhang, T. Y., Hao, Y. D., Xie, X. Q., Liu, X. W., ... & Lin, H. (2024). Predicting coronary heart disease in Chinese diabetics using machine learning. Computers in Biology and Medicine, 169, 107952.

[25] Saikumar, K., & Rajesh, V. (2024). A machine intelligence technique for predicting cardiovascular disease (CVD) using Radiology Dataset. International Journal of System Assurance Engineering and Management, 15(1), 135-151.

[26] Elmi, A. H., Abdullahi, A., & Barre, M. A. (2024). A machine learning approach to cardiovascular disease prediction with advanced feature selection. Indonesian Journal of Electrical Engineering and Computer Science, 33(2), 1-1x.

[27] Mijwil, M., Faieq, A. K., & Aljanabi, M. (2024). Early Detection of Cardiovascular Disease Utilizing Machine Learning Techniques: Evaluating the Predictive Capabilities of Seven Algorithms. Iraqi Journal For Computer Science and Mathematics, 5(1), 263-276.

[28] Verma, S., & Rubin, J. (2018, May). Fairness definitions explained. In Proceedings of the international workshop on software fairness (pp. 1-7).

[29] Feldman, M., Friedler, S. A., Moeller, J., Scheidegger, C., & Venkatasubramanian, S. (2015, August). Certifying and removing disparate impact. In proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining (pp. 259-268).

[30] Lee, N. T., Resnick, P., & Barton, G. (2019). Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. Brookings Institute: Washington, DC, USA, 2.
[31] Panch, T., Mattie, H., & Atun, R. (2019). Artificial intelligence and algorithmic bias: implications for health systems. Journal of global health, 9(2).
[32] Blattner, L., & Nelson, S. (2021). How costly is noise? Data and disparities in consumer credit. arXiv preprint arXiv:2105.07554.
[33] Cawley, G. C., & Talbot, N. L. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. The Journal of Machine Learning Research, 11, 2079-2107.
[34] Tantithamthavorn, C., McIntosh, S., Hassan, A. E., & Matsumoto, K. (2018). The impact of automated parameter optimization on defect prediction models. IEEE Transactions on Software Engineering, 45(7), 683-711.