International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Addressing Disparities in Cardiovascular Disease in Women and Minority Populations

Christiana Jemiyo [a], Doreen Ugwu [b], Adetola Anifat Ajayi [c] and Isaac Amornortey Yowetu [d]

[a] *Stellenbosch University, South Africa*
[b] *Jiann-Ping Hsu College of Public Health, Georgia Southern University, Georgia, USA*
[c] *University of West Georgia – Carrolton, GA*
[d] *Kwame Nkrumah University of Science and Technology, Ghana*

A B S T R A C T

Background: Cardiovascular disease (CVD) is a leading cause of morbidity and mortality, with women and minority populations experiencing a disproportionate burden. Despite advances in understanding CVD risk, disparities in its prevalence and outcomes remain, often influenced by a combination of social, biological, and behavioral factors. This study aims to identify the key factors contributing to these disparities and to better understand the specific risks faced by women and minority groups.

Method: We conducted an analysis using BRFSS, dataset to examine the impact of demographic, lifestyle, and health-related factors on CVD risk in women and minority populations. The study utilized generalized linear models (GLM) with logistic regression to assess the relationship between various predictors, including age, gender, race, smoking, exercise, alcohol consumption, and health history (flu, pneumonia) on CVD outcomes.

Results: Age was identified as the most significant predictor of CVD risk, with older individuals demonstrating a substantially higher likelihood of developing CVD. The interaction between gender and race revealed that minority women have an elevated risk compared to both male and majority counterparts. Lifestyle factors, including smoking and physical activity, were significantly associated with CVD outcomes. Smoking increased CVD risk, while regular exercise reduced it. Interestingly, alcohol consumption was associated with a 42% higher likelihood of CVD, challenging traditional assumptions about its protective effects. Additionally, a history of flu and pneumonia contributed to CVD risk, with pneumonia showing the strongest positive association.

Conclusion: The findings underscore the importance of addressing CVD risk disparities through targeted, age-specific, and culturally relevant interventions. Strategies should focus on promoting healthy behaviors, such as smoking cessation and regular physical activity, while considering the unique needs of women and minority populations. Future research should explore the mechanisms underlying these disparities and inform the development of more effective and equitable CVD prevention and management strategies.

Keywords: *Cardiovascular disease (CVD), Generalized Linear Model (GLM), Support Vector Machine (SVM), Logistic regression, Machine learning*

1. Introduction

Cardiovascular disease (CVD) remains the leading cause of death in the United States, posing a significant public health challenge that disproportionately impacts specific populations, particularly women and racial minorities. While it is often assumed that men are more susceptible to heart disease, women are equally at risk, albeit with CVD developing approximately 7 to 10 years later than in men. Despite this delayed onset, CVD remains the leading cause of death among women, a group often underrepresented in cardiovascular research. Misconceptions about hormonal protection and gender-specific symptoms contribute to delays in diagnosis and less aggressive treatment in women compared to men (Maas & Appelman, 2010).

Furthermore, racial and ethnic minorities, including African Americans and Hispanics, face compounded risks due to the prevalence of hypertension, diabetes, and limited access to quality healthcare. The intersection of these socio-economic and healthcare access barriers exacerbates the burden of cardiovascular disease in these populations. As the healthcare system continues to struggle with addressing these disparities, it is critical to identify the underlying factors contributing to the elevated risks and poor outcomes observed in these groups.

This research seeks to explore and address these disparities in cardiovascular health, specifically examining the unique risk factors, disease incidence, and treatment approaches for women and racial minorities. By analyzing large-scale datasets like the BRFSS (Behavioral Risk Factor Surveillance System), the study aims to develop strategies that can help mitigate these disparities through improved early detection, personalized prevention, and tailored treatment options for these populations.

2. Literature Review

The disproportionate impact of cardiovascular disease (CVD) on women and racial minorities has been well-documented in recent studies. While CVD is often associated with men, research has shown that women are significantly affected, with CVD being the leading cause of death in women (Maas & Appelman, 2010). Women, however, experience CVD at a later age than men—typically 7 to 10 years later. The assumption that women are inherently protected from CVD due to hormonal factors has contributed to a misunderstanding of their true risk. This misconception, coupled with the atypical symptoms women often present, results in delayed diagnoses and less aggressive treatment compared to men, contributing to worse health outcomes (Maas & Appelman, 2010).

In addition to gender, racial and ethnic minorities are also disproportionately affected by CVD. African Americans and Hispanics face a higher prevalence of CVD risk factors such as hypertension and diabetes, and these populations also face socio-economic barriers that limit their access to healthcare. Research by Ferdinand et al. (2020) emphasizes that Blacks are two to three times more likely to die from preventable heart disease and stroke compared to Whites. The lower rates of hypertension control and medication adherence, particularly among African Americans, are significant contributors to these disparities (Ferdinand et al., 2017). Systemic barriers such as economic strain, inadequate access to care, and chronic exposure to discrimination exacerbate the cardiovascular risk in these communities (Johnson et al., 2019).

Additionally, psychosocial factors play a significant role in CVD risk. Studies indicate that exposure to chronic stress and discrimination is correlated with higher rates of hypertension and CVD in minority populations. Johnson et al. (2019) underscores the importance of incorporating psychosocial support into cardiovascular care for communities experiencing systemic social stressors, which could offer a more holistic approach to reducing CVD risk.

While machine learning techniques have demonstrated significant potential in cardiovascular risk prediction, this study employs traditional statistical methods to analyze disparities in cardiovascular disease among women and minority populations. Nevertheless, reviewing machine learning advancements provides valuable context for understanding emerging tools and their potential application in future research.

Machine learning techniques have recently been applied to cardiovascular risk prediction with promising results. Studies by Pan et al. (2022), Alsaleh et al. (2023), and others have demonstrated that algorithms like Support Vector Machines (SVM), Random Forest (RF), and XGBoost can improve risk assessment by analyzing large datasets, providing more accurate early predictions, and allowing for tailored treatment plans. These approaches help in identifying significant risk factors for individuals, leading to proactive healthcare interventions. The use of SVM, in particular, has shown promise due to its effectiveness in high-dimensional datasets, minimizing generalization error and improving prediction accuracy (Cortes & Vapnik, 1995). Other models, such as hybrid methods that combine different algorithms, have been introduced to further enhance the accuracy and efficiency of cardiovascular disease predictions (Mohan et al., 2019; Ashri et al., 2021).

Despite these advancements, the integration of machine learning techniques in addressing healthcare disparities remains an emerging area of research. There is a need to further refine these models, especially with respect to minority populations, to improve the accuracy and inclusivity of predictive tools and healthcare interventions.

By examining these gender-specific and race-specific differences in risk factors, disease incidence, and treatment strategies, this literature highlights the need for more targeted, inclusive approaches in the prevention and treatment of cardiovascular disease. The development of machine learning models that can predict heart disease risk based on demographic, clinical, and psychosocial factors has the potential to enhance early intervention efforts and contribute to more equitable healthcare outcomes for women and racial minorities.

3. Materials and Methods

3.1. Data Collection and Integration

This study utilized cross-sectional data from the Behavioral Risk Factor Surveillance System (BRFSS), collected in 2020. The BRFSS is a nationwide, telephone-based survey conducted by the Centers for Disease Control and Prevention (CDC). It gathers comprehensive information on health-related risk behaviors, chronic health conditions, and the use of preventive services among adults in the United States. The dataset includes a wide range of variables such as demographics (age, gender, race, education level), socioeconomic factors (income, employment status), and self-reported health behaviors (physical activity, smoking, alcohol consumption, etc.). These data are critical for identifying health trends and disparities across different population groups.

3.2. Data Processing and Cleaning

To ensure the BRFSS dataset's suitability for this study, a thorough data cleaning and feature selection process was conducted. This involved addressing missing values, removing duplicate entries, and resolving inconsistencies that could introduce bias or inaccuracies in the analysis. Specifically, rows with missing data for key variables critical to the analysis, such as sex, health status, race, and education, were excluded.

The initial BRFSS dataset for 2020 contained 401,958 entries. Following the data cleaning process, 95,287 records which were retained for model development. These selected entries represent a high-quality subset of data, ensuring robust and reliable analysis.

3.3. Statistical Analysis

The core of the statistical analysis will focus on identifying both sex-specific and race-specific differences in cardiovascular disease (CVD) risk factors. A multivariable logistic regression will serve as the primary model to assess the influence of independent variables, such as age, race, gender, socioeconomic status, and health behaviors on CVD outcomes. This regression model will help quantify the likelihood of developing CVD based on these predictors, providing insight into which variables contribute most significantly to disparities in CVD risk.

To capture more nuanced relationships, interaction terms will be incorporated into the models. Specifically, gender and race interaction terms will be examined to determine whether certain combinations of risk factors disproportionately affect groups. For instance, it is hypothesized that the interaction of race and gender may amplify CVD risk for women in the minor (other races) compared to women in the majority group (Whites) in the United States. The analysis will also account for confounding factors such as lifestyle choices (e.g., smoking, alcohol, exercise) and age.

The initial dataset consisted of 401,958 entries before data cleaning. After removing incomplete or inconsistent data, the final study population comprises 95,287 individuals, divided into two primary age groups. Participants aged 45 years and older make up 53,211 individuals (approximately 55.8% of the total population), while those aged 18 to 44 years account for 42,076 individuals (approximately 44.2%).

In terms of gender distribution, the population includes 48,136 males (approximately 50.5%) and 47,151 females (approximately 49.5%), reflecting a nearly balanced representation of sexes. Regarding racial composition, the majority of the population identifies as White, with 67,889 individuals (approximately 71.2%), while minority racial groups account for 27,398 participants (approximately 28.8%).

Cardiovascular disease (CVD) is present in 3.5% of the population, with notable variations by demographic factors. Among individuals with CVD, 2,113 are males and 1,244 are females. Additionally, 2,711 cases of CVD are observed among Whites, while 646 cases occur in minority racial groups.

These descriptive statistics provide a robust demographic overview of the study population, offering valuable insights into analyzing disparities in CVD prevalence and associated risk factors. The data highlight important variations by age, gender, and race, which are essential for addressing health inequities and designing effective public health interventions.

3.4 Generalized Linear Model and Logistic Regression

Generalized Linear Models (GLMs) will be employed to model the relationships between CVD outcomes (binary dependent variables) and various predictor variables. Unlike traditional linear regression, which assumes normally distributed outcomes, GLMs are more flexible and can handle binary, count, and categorical outcome variables. For this study, logistic regression, a specific form of GLM, will be used to model the binary outcome of whether or not a person develops CVD.

Logistic regression is particularly suitable for predicting binary outcomes, such as the presence or absence of CVD, based on a combination of predictor variables. This method is ideal for analyzing how risk factors, such as age, gender, race, socioeconomic status, and behavioral patterns, impact the likelihood of developing CVD.

The logistic regression model will be specified as:

$$Log(p(y=1)/p(y=0)) = B0 + B1X1 + B2X2 + ... + BnXn$$

where p(y=1) represents the probability of an individual developing CVD, and X1, X2, ..., Xn of the independent variables (e.g., age, gender, race, cholesterol levels). The coefficients B1, B2, ..., Bn will be interpreted as log-odds ratios, providing insight into the relative contribution of each predictor to CVD risk.

This framework allows us to estimate the odds of developing CVD based on predictors such as age, gender, race, socioeconomic status, and lifestyle factors. Odds ratios will be used to compare the likelihood of CVD across different demographic groups, while confidence intervals will quantify the statistical uncertainty of these estimates.

Logistic regression is specifically used when the dependent variable (Y) is categorical and has two categories, typically coded as "1" to represent one category (e.g., the event of interest) and "0" to represent the other (e.g., the absence of the event) (Fritz and Berger, 2015). If the dependent variable has more than two categories, the analysis is referred to as multinomial logistic regression. However, this study will focus on binary logistic regression.

Additionally, the logistic regression model will be extended to include interaction terms to explore how the joint effect of race and gender influences CVD risk. This will provide a deeper understanding of intersectional disparities and identify high-risk subgroups that may require targeted interventions.

4. Results and Discussion

The study seeks to explore and address these disparities in cardiovascular health, specifically examining the unique risk factors, disease incidence, and treatment approaches for women and racial minorities.

4.1 Descriptive statistics on the data

The study provides insights into the prevalence of cardiovascular disease (CVD) across gender and racial groups. Out of the total male population (61,351), 2,685 men were reported to have CVD, representing approximately **4.4%** of the male population. Among females (60,733), 1,571 women were identified with CVD, accounting for approximately **2.6%** of the female population.

In terms of racial composition, among White individuals (88,106), 3,455 were reported to have CVD, which corresponds to approximately **3.9%** of the White population. Conversely, among minority racial groups (33,978), 801 individuals were found to have CVD, representing approximately **2.4%** of the minority population.

These statistics highlight disparities in CVD prevalence, with men and White individuals showing higher rates of CVD compared to women and minority racial groups. These findings are essential for tailoring public health interventions to address group-specific risk factors and improve outcomes.

Table 1: Descriptive Statistics of CVD Prevalence by Gender and Race

1. Category	**2. Subgroup**	**3. Total Population (N)**	**4. CVD Cases (n)**	**5. CVD Prevalence (%)**
6. Gender	7. Male	8. 48,136	9. 2,113	10. 4.39%
11.	12. Female	13. 47,151	14. 1,244	15. 2.64%
16. Race	17. White	18. 67,889	19. 2,711	20. 3.99%
21.	22. Minority	23. 27,398	24. 646	25. 2.36%
26. Total	27. All Individuals	28. 95,287	29. 3,357	30. 3.52%

This table summarizes the prevalence of cardiovascular disease (CVD) across gender and racial groups in the study population. Males and Whites exhibit higher CVD prevalence rates compared to females and minority racial groups.

4.2 Variance Inflation Factor (VIF) Analysis on the feature

Table 2: Summary on Variance Inflation Factor (VIF) on the feature

35. Features	**36. First Check of VIF**	**37. Second check of VIF**
38. Education	39. 24.821713	40. ---
41. Sex	42. 2.038161	43. 1.949407
44. Age	45. 2.536882	46. 2.301504
47. Race	48. 1.500078	49. 1.422634
50. Income	51. 11.094523	52. ---
53. Smoker	54. 1.229178	55. 1.178059
56. Diabetes	57. 9.130261	58. ---
59. HIV	60. 11.513485	61. ---
62. Alcohol	63. 2.409290	64. 5.121664
65. Flu	66. 2.024158	67. 1.921529
68. Pneumonia	69. 1.507450	70. 1.496147
71. BMI	72. 15.395447	73. ---
74. Exercise	75. 5.063257	76. 3.321681

Multicollinearity was assessed using Variance Inflation Factor (VIF) analysis. Initial analysis revealed high VIF values for features such as Educ (24.82), Income (11.09), HIV (11.51), BMI (15.40), and Diabetes (9.13), indicating significant multicollinearity. These features were excluded to enhance model stability.

The refined model retained predictors with acceptable VIF values, all below 5, including Sex, Age, Race, Smoker, Alcohol, Exercise, Flu, and Pneumonia. This adjustment ensures reliable coefficient estimates and robust conclusions regarding the impact of these predictors on CVD outcomes.

Table 3: a. Results of the Generalized Linear Model

77. Dep. Variable:	78. CVD	79.	80. No. Observation:	81. 95287
82. Model:	83. GLM	84.	85. Df. Residuals:	86. 95277
87. Model Family:	88. Binomial	89.	90. Df. Model:	91. 9
92. Link Function:	93. Logit	94.	95. Scale	96. 1.0000
97. Method:	98. IRLS	99.	100. Log-Likelihood:	101. -12644
102. Date	103. Dec. 02, 2024	104.	105. Deviance:	106. 25287
107. No. Iterations:	108. 8	109.	110. Pearson Chi2:	111. 1.02e+05
112. Covariance Type:	113. nonrobust	114.	115. Pseudo R-Squ. (CS):	116. 0.03880

b. Results of the Generalized Linear Model

117.	118. Coef	119. Std err	120. z	121. P>|z|	122. [0.025	123. 0.975]
124. Constant	125. -5.2546	126. 0.093	127. -56.209	128. < 0.001	129. -5.438	130. -5.071
131. Age	132. 1.9211	133. 0.065	134. 29.777	135. < 0.001	136. 1.795	137. 2.048
138. Sex	139. -0.7833	140. 0.042	141. -18.564	142. < 0.001	143. -0.866	144. -0.701
145. Race	146. -0.3867	147. 0.061	148. -6.304	149. < 0.001	150. -0.507	151. -0.266
152. SRTerm	153. 0.4695	154. 0.092	155. 5.130	156. < 0.001	157. 0.290	158. 0.649
159. Smoker	160. 0.2850	161. 0.046	162. 6.141	163. 0.001	164. 0.194	165. 0.376
166. Alcohol	167. 0.3482	168. 0.037	169. 9.479	170. < 0.001	171. 0.276	172. 0.420
173. Flu	174. 0.2140	175. 0.039	176. 5.459	177. < 0.001	178. 0.137	179. 0.291
180. Pneumonia	181. 1.0752	182. 0.039	183. 27.720	184. < 0.001	185. 0.999	186. 1.151
187. Exercise	188. -0.3621	189. 0.040	190. -8.988	191. < 0.001	192. -0.441	193. -0.283

4.3 Generalized Linear Model Regression Results

The results of the Generalized Linear Model (GLM) regression analysis provide valuable insights into the predictors of cardiovascular disease (CVD). The analysis included 95,287 observations and employed a binomial family with a logit link function to examine the impact of various factors on the odds of developing CVD. The model demonstrates a pseudo R-squared value of 0.03880, highlighting its explanatory power. The regression results are as follows:

- Age: Age is a strong predictor of CVD. The exponential coefficient exp(1.9211)=6.83 indicates that older individuals have approximately 6.83 times higher odds of developing CVD compared to younger individuals.
- Sex: Females have significantly lower odds of CVD compared to males. The odds ratio exp(−0.7833) = 0.46 indicates a 54% reduction in risk for females.
- Race: Minority racial groups show reduced odds of CVD compared to the majority (Whites). The odds ratio exp(−0.3867) =0.68 suggests a 32% decrease in risk for minority racial groups.
- Interaction Between Sex and Race (SRTerm, B=0.4695, p<0.001): The interaction term between Sex and Race significantly impacts CVD odds. The odds ratio exp(0.4695)=1.60 suggests that minority females are 60% more likely to develop CVD compared to majority males, reflecting the intersectional nature of demographic predictors.
- Smoking: Smoking is associated with a significant increase in CVD risk, with an odds ratio exp(0.2850)=1.33, representing a 33% increased likelihood of CVD.
- Alcohol Consumption (B=0.3482, p<0.001): Alcohol use increases the odds of CVD. The exponential coefficient exp(0.3482)=1.42 indicates a 42% greater likelihood of CVD among alcohol consumers.
- Influenza: A history of influenza is associated with higher odds of CVD. The odds ratio exp(0.2140) =1.24 reflects a 24% increased risk.
- Pneumonia: Pneumonia is a substantial predictor of CVD. The odds ratio exp(1.0752) = 2.93 suggests that individuals with pneumonia are nearly three times more likely to develop CVD.

- Exercise: Regular physical activity significantly reduces CVD risk. The odds ratio $\exp(-0.3621) = 0.70$ corresponds to a 30% reduction in odds for individuals who engage in exercise.

The GLM results explain the importance of both demographic and behavioral factors in predicting CVD risk. Age emerges as the most substantial risk factor, while modifiable behaviors such as smoking, alcohol consumption, and physical inactivity provide opportunities for intervention.

The significant interaction between Sex and Race highlights the complexity of health disparities, with minority females facing heightened vulnerability to CVD. This underscores the necessity of culturally sensitive and targeted interventions to address these inequities effectively.

Our findings on age as a dominant risk factor are consistent with Maas & Appelman (2010), who identified advancing age as a major determinant of CVD risk. Similarly, the reduced odds of CVD in females align with previous studies attributing this trend to hormonal differences and delayed onset of disease. However, our observation of lower odds of CVD in minority racial groups contrasts with Ferdinand et al. (2020), who reported higher CVD prevalence in these populations. This discrepancy may be attributable to differences in population demographics or access to healthcare services in our study sample. Behavioral factors like smoking and alcohol consumption, identified as significant predictors in our study, reinforce findings from public health literature on modifiable risk factors. Moreover, our identification of a significant interaction between sex and race highlights a unique dimension of health disparities, underscoring the importance of intersectional approaches in CVD research and intervention strategies.

5. Conclusion

This study identifies critical factors influencing cardiovascular disease (CVD) risk, with age emerging as the most significant predictor. Older individuals exhibit a significantly higher likelihood of developing CVD, emphasizing the need for age-specific preventive strategies and interventions. Demographic factors, such as gender and race, also play significant roles. While males and Whites exhibit higher baseline risks, the interaction between gender and race reveals an elevated risk for minority females, highlighting the intersection of social and biological determinants.

However, the study has limitations, including reliance on self-reported data, and geographic variability in survey responses, which may impact the generalizability of the findings. Future research should focus on understanding the mechanisms behind these relationships and addressing these limitations to improve CVD prevention strategies.

Lifestyle factors are critical in influencing CVD outcomes. Smoking is associated with an increased risk, while regular exercise serves as a protective factor. Alcohol consumption also contributes to an increased risk of CVD disease. Health-related factors, including a history of flu and pneumonia, are also significant, with pneumonia showing the strongest positive association with CVD risk.

These findings emphasize the importance of comprehensive, targeted interventions to reduce CVD disparities. Strategies should focus on promoting healthy lifestyles, including smoking cessation, regular exercise, and reducing alcohol consumption, while addressing structural barriers to healthcare access. Future research should explore the mechanisms underlying the gender-race interaction and other predictors to refine prevention strategies and inform equitable health policies.

Disclaimer

The views expressed in this article are those of the authors and do not necessarily represent the official policies or positions of any affiliated institutions.

Authorship

All authors contributed equally to the conceptualization, writing, and editing of this article without any external writing assistance.

Conflict of Interest

The authors declare no conflicts of interest related to this research.

References

- Ali, M.M., Paul, B.K., Ahmed, K., Bui, F.M., Quinn, J.M., Moni, M.A., 2021. Heart disease prediction using super- vised machine learning algorithms: Performance analysis and comparison. Computers in Biology and Medicine 136, 104672.
- Alsaleh, M.M., Allery, F., Choi, J.W., Hama, T., McQuillin, A., Wu, H., Thygesen, J.H., 2023. Prediction of disease comorbidity using explainable artificial intelligence and machine learning techniques: A systematic review. Interna- tional Journal of Medical Informatics 175, 105088.
- Ashri, S.E., El-Gayar, M.M., El-Daydamony, E.M., 2021. Hdpf: heart disease prediction framework based on hybrid classifiers and genetic algorithm. ieee access 9, 146797–146809.
- Carnethon, M.R., Pu, J., Howard, G., Albert, M.A., Anderson, C.A., Bertoni, A.G., Mujahid, M.S., Palaniappan, L., Taylor Jr, H.A., Willis, M., et al., 2017. Cardiovascular health in african americans: a scientific statement from the american heart association. Circulation 136, e393–e423.
- Cortes, Corinna. "Support-Vector Networks." *Machine Learning* (1995).

- Duraisamy, B., Sunku, R., Selvaraj, K., Pilla, V.V.R., Sanikala, M., 2024. Heart disease prediction using support vector machine. Multidisciplinary Science Journal 6.
- Ferdinand, K., Batieste, T., Fleurestil, M., 2020. Contemporary and future concepts on hypertension in african americans: Covid-19 and beyond. Journal of the National Medical Association 112, 315–323.
- Ferdinand, K.C., Yadav, K., Nasser, S.A., Clayton-Jeter, H.D., Lewin, J., Cryer, D.R., Senatore, F.F., 2017. Disparities in hypertension and cardiovascular disease in blacks: the critical role of medication adherence. The Journal of Clinical Hypertension 19, 1015–1024.
- Fitriyani, N.L., Syafrudin, M., Alfian, G., Rhee, J., 2020. Hdpm: an effective heart disease prediction model for a clinical decision support system. IEEE Access 8, 133034–133050.
- Fritz, M., Berger, P.D., 2015. Improving the user experience through practical data analytics: gain meaningful insight and increase your bottom line. Morgan Kaufmann.
- Hsu, Chih-Wei. "A Practical Guide to Support Vector Classification." *Department of Computer Science, National Taiwan University* (2003).
- Johnson, D.A., Jackson, C.L., Williams, N.J., Alca´ntara, C., 2019. Are sleep patterns influenced by race/ethnicity–a marker of relative advantage or disadvantage? evidence to date. Nature and science of sleep , 79–95.
- Li, J.P., Haq, A.U., Din, S.U., Khan, J., Khan, A., Saboor, A., 2020. Heart disease identification method using machine learning classification in e-healthcare. IEEE access 8, 107562–107582.
- Maas, A.H., Appelman, Y.E., 2010. Gender differences in coronary heart disease. Netherlands Heart Journal 18, 598– 603.
- Mohan, S., Thirumalai, C., Srivastava, G., 2019. Effective heart disease prediction using hybrid machine learning techniques. IEEE access 7, 81542–81554.
- Mosca, L., Benjamin, E.J., Berra, K., Bezanson, J.L., Dolor, R.J., Lloyd-Jones, D.M., Newby, L.K., Pin˜a, I.L., Roger, V.L., Shaw, L.J., et al., 2011. Effectiveness-based guidelines for the prevention of cardiovascular disease in women—2011 update: a guideline from the american heart association. Journal of the American College of Cardi- ology 57, 1404–1423.
- Ouyang, S., 2022. Research of heart disease prediction based on machine learning, in: 2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), IEEE. pp. 315–319.
- 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.
- Rahim, A., Rasheed, Y., Azam, F., Anwar, M.W., Rahim, M.A., Muzaffar, A.W., 2021. An integrated machine learning framework for effective prediction of cardiovascular diseases. IEEE Access 9, 106575–106588.