INVESTIGATING GENDER BIAS IN THE DIAGNOSIS OF HEART DISEASE

UZORDINMA, GRACE (Q2241217)

GENDER BIAS IN PREDICTING HEART DISEASE

School of Computing and Digital Technologies, Teesside University, United Kingdom

Artificial Intelligence Ethics and Applications - CIS4057-N-FJ1-2023

ABSTRACT

In this research, I investigated the impact of patient gender bias on the likelihood of seeing a cardiologist and been diagnosed with heart failure. I draw the conclusion that there is gender bias in both access and diagnosis for patients who have a high risk of developing heart problems using data from Kaggle. After adjusting for risk factors and demographic features, my results indicate that there is no equal opportunity, group fairness and accuracy between the female group and male group. This might also be because women are less likely to visit a cardiologist and to receive a heart disease diagnosis according to O. Alonso Gelabert et al.

1. INTROUCTION

Women are still frequently misdiagnosed or underdiagnosed with cardiac problems. Women are 50% more likely to be given an incorrect initial diagnosis, which can be fatal while they are experiencing a heart attack.

According to the World Health Organisation (WHO), one of the main causes of death worldwide is cardiovascular diseases (CVD), which accounts for 35% of all deaths worldwide in 2019, women have an even higher relative risk of CVD morbidity and mortality. [S. Mendis, et al 2011].

Seeing as CVD is more common in men, it has historically been described to as a "male" disease, and medical research on this condition have frequently involved male patients. As a result, symptoms including cold chills, nausea, vomiting, and peculiar fatigue that are more common in women are not recognised Vogel, B, et al 2021).

My goal is to find out if there is any form of discrimination against women in the likelihood of seeing a cardiologist or being diagnosed with a cardiac problem. This study highlights the significance of resolving any biases that may exist within the healthcare system and is a critical step in understanding the underlying causes of discrepancies in healthcare outcomes and access. At the end of this work, I would advise that policies and practices that support equitable access to healthcare for all people should be put in place, regardless of gender, by offering insight into this understudied field. [O. Alonso, et al. 2023]

Though physiological and pathological differences between men and women have been observed, such as narrower arteries, different electrical properties, and different plaque composition and development in women, the main risk factors for heart disease were previously thought to be the same for both genders [M¨oller-Leimku¨hler, et al. 2007]. Legato [M.J. Legato, 1997] suggests that these variations may lead to different CVD patterns in men and women.

2. LITERATURE REVIEW

Märit Mejhert, et al (2021) compared to know if there is gender bias in the coherence to guidelines in the diagnostic test, treatment and follow-up in heart failure patient. The study highlights inconsistencies in therapy, follow-up, and diagnostic test usage for heart failure patients, particularly in women, requiring careful interpretation due to retrospective results.

O. Alonso Gelabert, et al (2023). Examined the gender bias in the diagnosis of cardiovascular disorders in Catalonia. This study reveals gender bias in access and diagnosis of heart disease, with women having lower probabilities of visiting a cardiologist and being diagnosed, despite controlling for risk factors and demographic characteristics.

3. ETHICAL CONCERNS

Data Security and Privacy: The management of private and sensitive medical data poses issues with data security and privacy, making the use of encryption, anonymization, and secure storage necessary to protect personal information.

Bias and Fairness: AI algorithms should be trained with fairness and accuracy, ensuring that predictions are fair and equitable for all individuals, regardless of race, gender, or socioeconomic status.

Informed Consent: Patients should be given informed consent before using their medical data in AI models, they should be informed about the use of their data, potential risks and benefits, and their right to opt-out.

Algorithmic governance: In order to prevent the misuse or abuse of AI technologies in predicting heart disease and to encourage responsible AI creation, ethical principles, standards, and laws should be put in place.

4. METHODOLOGY

4.1. DATASET DESCRIPTION

This dataset was gotten from Kaggle with 918 observations and 12 attributes.

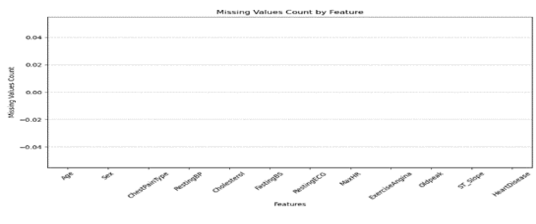
https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction/data

All 11 attributes are diagnostic test that need to be carried out in order to determine if a patient has a heart disease. The table below shows all the attributes with a description of each for better understanding of what I am working with.

4.2. DATA PRE-PROCESSING

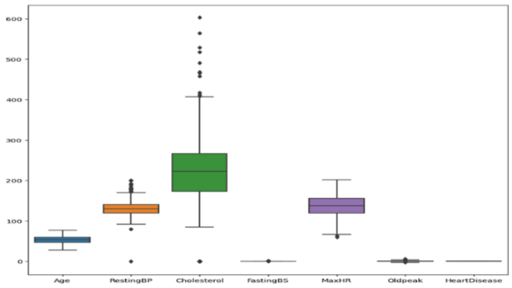
4.2.1. Missing Data:

I checked for missing or null values, but my code showed false, that means I do not have a missing value in my data.



4.2.2. Handling Outliers:

I used the box plot to check for outliers, though I found that they were outliers however they are not significant so therefore, I did not deal with it.



4.2.3. Removing Duplicate:

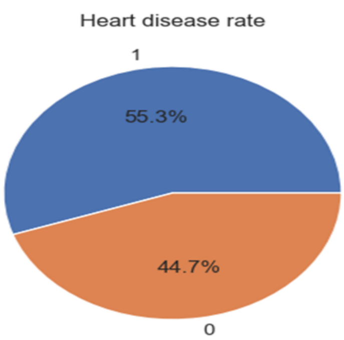
Using the duplicate function, I realized I do not have duplicates in my dataset.

```
0      False
1      False
2      False
3      False
4      False
       ...
913    False
914    False
915    False
916    False
917    False
Length: 918, dtype: bool
```

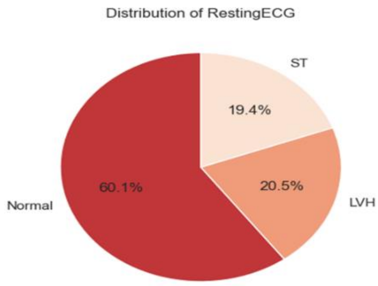
4.3. EXPLORATORY DATA ANALYSIS

4.3.1. Uni-variate Analysis

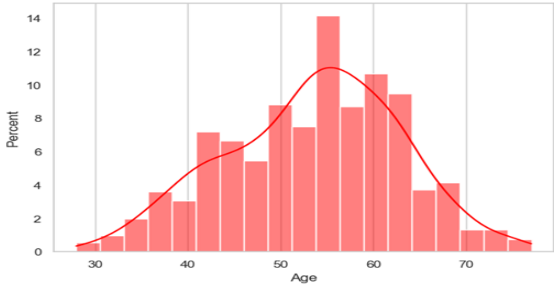
Heart disease Rate chart: shows that 55% of the patient in this dataset has heart disease while 44% do not have heart failure, hence when running the machine learning model there will be no need to up sample the target variable as there will be no biased accuracy as regards this.



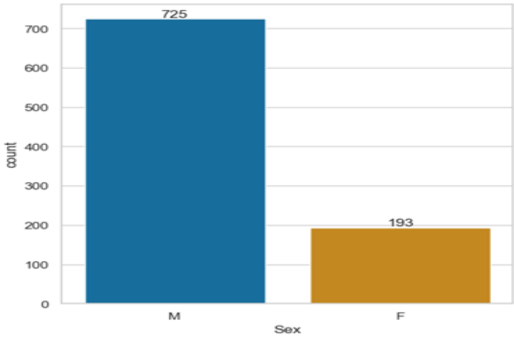
Distribution of RestingEcg chart: Resting electrocardiogram readings are referred to as restingECG. The terms "Normal" indicates a normal result, "ST" denote the presence of abnormalities, and "LVH" shows left ventricular hypertrophy. The pie chart shows that 60.1% has normal result, 20.5% has LVH and 19.4% shows ST abnormalities.



Age Distribution Chart: This chart shows that people around the age of 40 to 70 visit the cardiovascular hospital to check if they have a heart disease or going for treatment. However, it is advisable that people within the age of 20-30 and above 70 should also visit occasionally just to check that they are healthy.

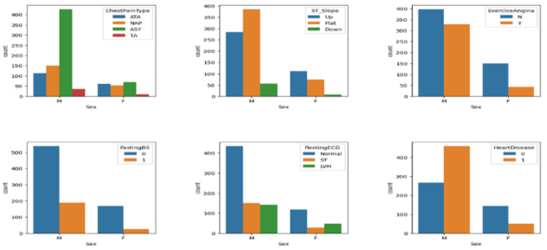


Gender Distribution: From this chart, it is observed that most of the observation are the male with 725 observation and the female with 193. which literally shows that there is bias in the data.

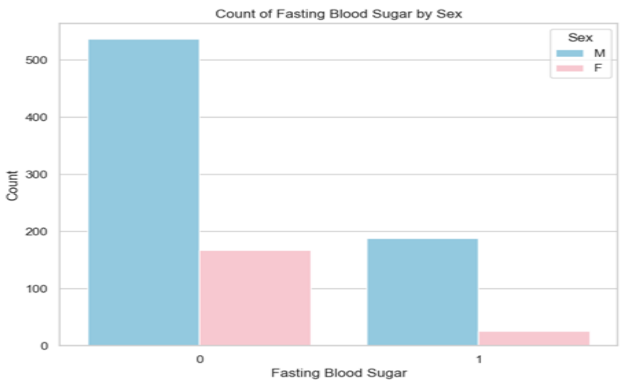


4.3.2. Bi-variate Analysis:

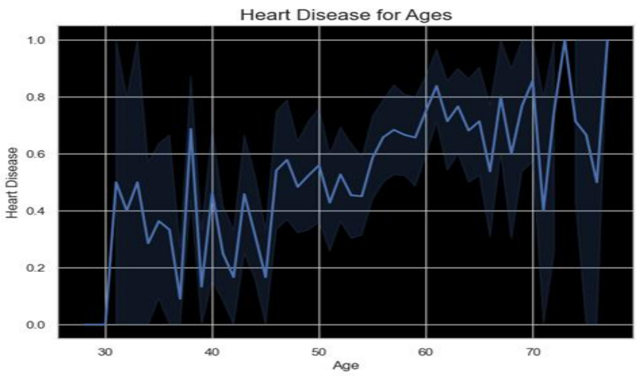
: Fig 1. shows a count plot of all the categorical variables, this also shows that for all the diagnostic test the male gender is always topping the chart which depicts that there are more male than female that visited the cardiology clinic with an obvious bias in the dataset.



Fasting Blood sugar by Gender: The plot shows that there are greater number of people with low blood sugar. the chart also shows more male with high FBS than female. this can also be seen as a bias because the dataset has more male than the female and the general opinion of people that the men then to have heart failure than the women.

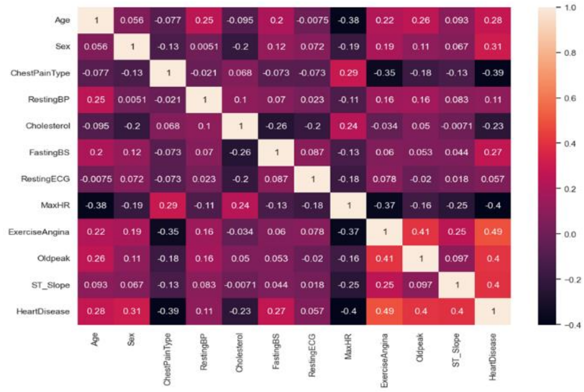


Heart Disease for the Age distribution: This shows a line chart of people with heart disease in relation to their age. and here I can see that people below the age of 30 are likely to not have a heart failure and people above 70 are more likely to have heart failure.



4.4. **Correlation Analysis**:

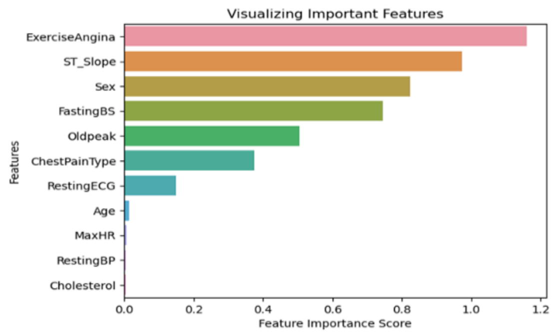
It is a statistical method for determining the direction and intensity of a relationship between two. Strong positive relationships are shown by correlation coefficients close to +1, strong negative relationships are indicated by coefficients close to -1, and little to no relationships are indicated by coefficients close to 0 between the variables. From my dataset, it shows that ExceriseAngina, Oldpeak, ST-Slope are a little correlated to my target variable HeartDisease.



5. MACHINE LEARNING MODEL

5.1. Feature Selection:

The SVM algorithm is used for feature selection to achieve better predictive accuracy and features are selected from a model in the scikit-learn library. the feature selection showed that the RestingBp, Cholesterol, MaxHR will not perform better in the model.



5.2. Proposed Algorithms:

The SVM was utilized for data handling and checking high dimensionality, while KNN was employed for neighbourhood selection and for comparing accuracy with SVM.

5.3. Evaluation Process Used:

The evaluation process utilizes, accuracy score, precision, recall, F1 score a confusion matrix, which is a table-like structure with true positive and false negative values.

Accuracy = (TP+TN)/(TP+TN+FP+FN)

precision = (TP)/(TP+FP)

Recall = (TP)/(TP+FN)

Positive Rate= (TP+ FP) / (TP+ FP + TN + FN)

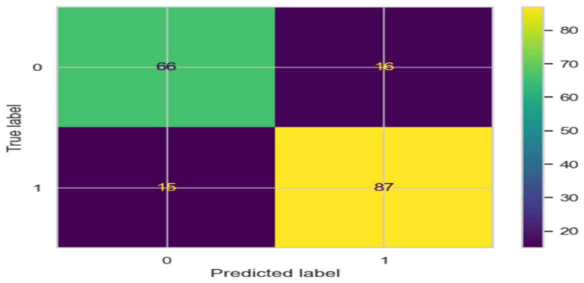
5.4. Splitting of Dataset:

After the data preprocessing and EDA, I split my dataset into 70% training (734 rows, 11 columns) and 30% testing (184 rows, 11 columns) afterwards, I went ahead to standardize the features eliminates biases, equalises scales, and satisfies algorithmic assumptions, all of which contribute to better machine learning model performance and interpretability.

5.5. SVM Model and Result:

The model achieved high accuracy, precision, recall, and specificity, with 83.15% of predictions being correct. It correctly identified 85.29% of positive cases and 80.49% of negative cases. The model performed reasonably well in making predictions, indicating high accuracy, precision, recall, and specificity.

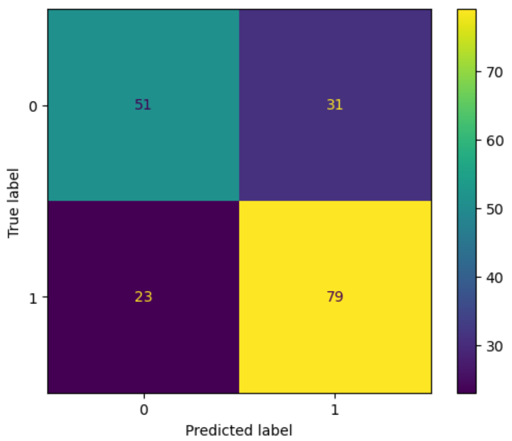
Confusion Matrix:



5.6. KNN Model and Result:

I ran this model in order to compare with the SVM to know which model will do better as regards the prediction of heart disease. the result showed that SVM did better than KNN with the below result. The model achieved 70.65% accuracy, 71.82% precision, and 77.45% recall on testing data, indicating it correctly classified 70.65% instances, 71.82% of true positive predictions, and 77.45% of actual positive cases, indicating decent performance but room for improvement.

Confusion Matrix:



5.7. Results

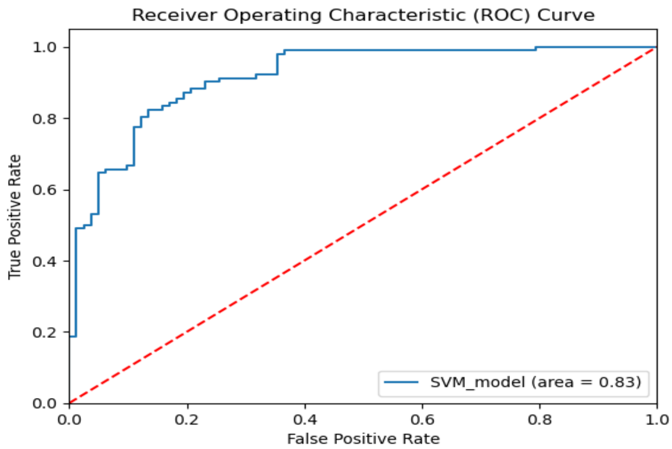
Model	Accuracy	precision	recall	f1-score
SVM	83%	84%	85%	85%
KNN	71%	72%	77%	75%

5.8. Hyperparameter tunning:

I carried out this tunning in order to optimize the performance of my model by finding the best combinations of hyperparameters. and I realize that that the model accuracy improved after optimizing the SVM model improved in its accuracy from an 83% to 86% and the KNN model improved from 70% to 75%.

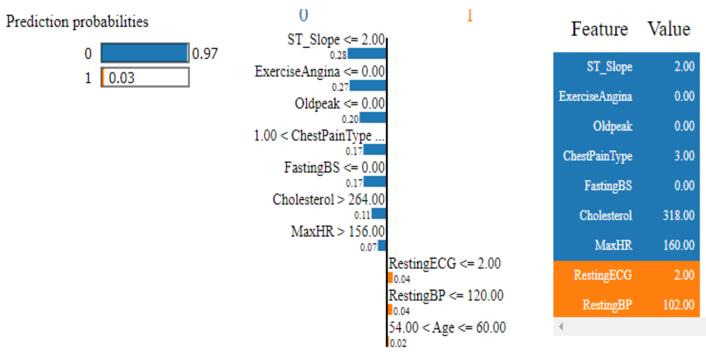
ROC CURVE:

The ROC curve visually represents the model's performance, assessing its discriminative ability. It shows true positive rate and false positive rate as classification threshold changes, with a closer curve indicating better performance.



5.9. Explainable AI Using LIME:

The Lime explanation offers transparency into the model's decision-making process, enabling stakeholders to assess its behaviour and identify potential biases, enhancing its trustworthiness and real-world application deployment. This study reveals that features like ST_Slope, ExerciseAngina, and Oldpeak have a significant negative impact on heart disease risk, while RestingECG and RestingBP have a positive impact.



5.10. Fairness evaluation

Equal Accuracy: This is the same accuracy across groups and the same percentage of correct classification in groups.

	Male	
	PP	PN
P	TP (85)	FN (9)
N	FP (14)	TN (38)

Accuracy= 123/146=84%

	Female	
	PP	PN
P	TP (2)	FN (6)
N	FP (2)	TN (28)

Accuracy=30/38=78%

Group Fairness: The criterion ensures model predictions are consistent across different demographic groups, ensuring positive predictions are similar across all demographic groups.

	Male	
	PP	PN
P	TP (85)	FN (9)
N	FP (14)	TN (38)

Positive Rate= 99/146=67%

	Female	
	PP	PN
P	TP (2)	FN (6)
N	FP (2)	TN (28)

Positive Rate= 4/38=11%

Equal Opportunity: This ensures equitable predictive performance across demographic groups by achieving parity in true positive and false positive rates, minimizing disparities in false positives and false negatives.

	Male	
	PP	PN
P	TP (85)	FN (9)
N	FP (14)	TN (38)

Recall= 85/94=90%

	Female	
	PP	PN
P	TP (2)	FN (6)
N	FP (2)	TN (28)

Recall= 2/8=25%

6. FINDING AND DISCUSSION

The machine learning model was carried out in order to see how well a gender-based predictive model for heart disease diagnosis performs. Despite the model's overall high accuracy and precision, a closer look reveals possible disparities and bias between gender in terms of recall and positive rate.

SVM and KNN Model: The prediction of heart disease was rather accurate for both the SVM and KNN models. However, the accuracy, precision, and recall of the SVM model were marginally higher than those of the KNN model, with the accuracy of SVM been 83% and KNN been 70%

Hyperparameter Tunning: Hyperparameter tuning improved SVM performance to 86%, with linear kernel and regularization parameters enhancing accuracy, while Manhattan distance metric and neighbourhood weights improved KNN's performance to 75%.

Protected Characteristics: The high percentage in recall (90%) for male patients indicates that the model is more accurate in identifying cases of heart disease in male patients than in female patients. This suggests that the model's capacity to identify positive cases may be biased towards women and there is no equal opportunity between the gender groups. In addition, the positive rate for men is significantly higher than for women with male having 67% and the women with 11%, suggesting that the model favours predicting heart disease results for men regardless of the patients' true conditions and there is no fairness among both genders. Finally, the percentage of accuracy for the male patients is 84% which is higher than the female patient and this suggests that there is no equal and correct classification for both genders. The accuracy of the results may seem satisfactory, but further examination is necessary to identify potential biases.

7. Conclusion

Gender discrimination is a multifaceted issue, influenced by gender-specific characteristics and persistent societal stereotypes, affecting various fields beyond healthcare (O. Alonso Gelabert, et al.). The differences and bias from my analysis may raise questions about how equitable and fair the predictive model is for different gender groupings. It is imperative to tackle these discrepancies in order to guarantee the validity and impartiality of the model's forecasts for every patient, regardless of gender. To address these concerns about fairness as well as enhance the model's performance across a range of demographic groups, more research and even a re-evaluation of the model's features and training data may be required.

8. REFERENCES

[1]. Mendis, S., Puska, P., & Norrving, B. (2011). Global atlas on cardiovascular disease prevention and control. World Health Organization, 2-14.

[2]. Gelabert, O. A., Veciana, M. B., Valsells, V. B., Baqe, M. C., & Nicodemo, C. (2023). Gender bias in the diagnosis of cardiovascular disorders in Catalonia. Health Policy, 132, 104823. https://doi.org/10.1016/j.healthpol.2023.104823

[3]. Moller-Leimkuhler, A. M. (2007). Gender differences in cardiovascular disease and comorbid depression. Dialogues in Clinical Neuroscience, 9(1), 71-83. https://doi.org/10.31887/dcns.2007.9.1/ammoeller

[4]. Legato, M. J. (1997). Gender-specific physiology: how real is it? How important is it? International Journal of Fertility and Women's Medicine, 42, 19-29.

[5]. Mejhert M, Holmgren J, Wändell P, Persson H, Edner M. Diagnostic tests, treatment and follow-up in heart failure patients--is there a gender bias in the coherence to guidelines? Eur J Heart Fail. 1999 Dec;1(4):407-10. doi: 10.1016/s1388-9842(99)00053-7. PMID: 10937955.

[6]. Vogel, B., Acevedo, M., Appelman, Y., Bairey Merz, C.N., Chieffo, A., Figtree, G.A., & Mehran, R. (2021). The Lancet women and cardiovascular disease Commission: reducing the global burden by 2030. The Lancet.