Available online at www.SciMedJournal.org

SciMedicine Journal

(ISSN: 2704-9833)

Vol. 5, No. 2, June, 2023

Comparative Study on Prediction of Survival Event of Heart Failure Patients Using Machine Learning and Statistical Algorithms

O. E. Oyewunmi [1], O. B. Aladeniyi [1], O. K. Bodunwa [1*]

[1] *Federal University of Technology, Akure, Nigeria.*

Received 09 February 2023; Revised 21 May 2023; Accepted 28 May 2023; Published 01 June 2023

Abstract

In a pressing global health concern with substantial morbidity and mortality rates, accurate survival prediction is paramount for informed decision-making and enhanced patient well-being. This study presented a comparative investigation aimed at predicting the survival events of heart failure (HF) patients through the utilization of both machine learning and statistical algorithms. A comprehensive dataset drawn from Allied Hospital and the Faisalabad Institute of Cardiology, Faisalabad, Pakistan, was used. The Synthetic Minority Over-Sampling Technique (SMOTE) was employed on the data to rectify the imbalance, and a notable improvement was observed. To ascertain significant variables, statistical methods (Mann-Whitney and Chi-Square) were compared with machine learning-based feature selection to identify pivotal features for survival prediction, namely ejection fraction and serum creatinine. Remarkably, on final training with these features, the Random Forest Classifier emerges as the top-performing model, boasting an accuracy exceeding 90%. These findings hold the potential to substantially enhance patient prognosis, management, and outcomes, consequently alleviating the strain on healthcare systems.

Keywords: Survival Prediction; Heart Failure; Machine Learning; Statistical Algorithms; Robust Predictor.

1. Introduction

Heart failure (HF) is a condition in which the heart cannot pump adequate blood to the body. Coronary heart disease, excessive blood pressure, and other chronic heart illnesses or diseases are the main causes [1]. According to data from the WHO, HF is one of the leading sources of mortality, taking the lives of around 17.5 million people each year, or roughly 32% of all fatalities worldwide. According to the Centers for Disease Control and Prevention, Americans experience a heart attack every single 40 seconds [2–4]. The danger of HF in patients can be determined by means of a variety of approaches, including biomarkers, risk scores, and their combination [5]. However, the predictive effectiveness of traditional risk prediction methods for HF is relative. Given the significance of an important organ like the heart, physicians have made it a priority to anticipate HF; nevertheless, to date, clinical practice has not consistently been able to foresee HF-related events with high accuracy [6, 7].

Understanding future risks and expected disease trajectories is critical for patient-physician interactions, allowing up-to-date decisions about treatment plans or end-of-life care [8]. Identifying low-risk patients can also

---

* Corresponding author: okbodunwa@futa.edu.ng

http://dx.doi.org/10.28991/SciMedJ-2023-05-02-01

help reduce patient tension and needless treatment costs. However, determining a patient's risk of HF presents challenges, as the complexity of HF produces vast amounts of data that can devastate clinicians, requiring simultaneous assessment of multiple components and their relationships. To address this issue, a consistent decision support system based on artificial intelligence and machine learning techniques can aid clinicians in analyzing patient records and making informed decisions. Traditional statistics and artificial intelligence groups gave rise to machine learning [9, 10].

Machine learning algorithms have shown capacity for predicting survival events for HF patients due to their emphasis on predictive performance and model generalization, which traditional statistics may not be able to carry out effectively [11, 12]. Given the extensive data acquired, machine learning approaches offer a practicable solution for prediction in HF. According to the study of Chicco & Jurman (2020) on 299 patients with heart failure, their data comprised 13 characteristics, such as smoking, sex, and high blood pressure. To forecast the percentage of survivors and rank the features related to the most significant risk variables, they applied univariate statistical analysis to rank the most relevant factors contributing to a higher risk of death in heart failure patients. They provided empirical evidence that only two features, serum creatinine and ejection fraction, were sufficient to predict the survival of heart failure patients [7]. Also, to indicate the mortality of HF patients hospitalized at the Institute of Cardiology and Allied Hospital in Faisalabad, Pakistan [13], traditional biostatistics time-dependent models, such as Cox regression and Kaplan-Meier survival plots, were used. Advancing age, renal dysfunction, blood pressure, anemia, and ejection fraction were recognized as significant risk factors. Two distinct gender-based survival prediction models were put forth in the same study. They stated that there were considerable differences between the main risk variables and survival prediction models for heart failure patients who were male and female. However, this claim must be assessed in a broader population to generalize the findings [14, 15]. Some other literature reviewed in the process of this study includes, but is not limited to, the following: Edoctor: Machine Learning and the Future of Medicine [16]; Identifying prognostic features for predicting heart failure by using machine learning algorithms [17]; Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging [18]; A primer in Artificial Intelligence in cardiovascular medicine [19] supervised deep learning embeddings for the prediction of cervical cancer diagnosis [20]; Heart disease and stroke statistics—2021 update [21].

Hence, there is a need for an accurate prediction of the survival rate of HF patients that would provide healthcare providers with appreciated information for informed decision-making and enhance patient outcomes. By understanding the necessary factors that contribute to the survival prediction of HF patients, healthcare providers can prioritize the factors that are most significant in the treatment and management of these patients. This study, therefore, develops a precise survival prediction model for HF patients by relating various machine learning techniques and classifying essential variables impacting survival prediction. The best model would be instrumental in improving prognosis and management, leading to better patient outcomes and reduced morbidity and mortality rates associated with HF.

2. Materials and Methods

2.1. Dataset Description

The dataset used in this study comprises medical records of 299 heart failure patients from the Faisalabad Institute of Cardiology and the Allied Hospital in Faisalabad, Pakistan, in 2015 [22]. Patients with left ventricular systolic dysfunction were classified as class III or IV of the New York Heart Association classification of the stages of heart failure. The dataset contains 11 clinical features, a follow-up period variable, and a binary response variable indicating patient survival. Features include age, anemia, high blood pressure, creatinine phosphokinase (CPK) levels, diabetes, ejection fraction, sex, platelets, serum creatinine, serum sodium, smoking, and follow-up period. The dataset was pre-processed, and the follow-up period variable was excluded from the study to focus on clinical feature-based patient risk identification. Data available at the UCI Machine Learning Repository: Heart Failure Clinical Records Data Set [23].

2.2. Data Preprocessing

The dataset was pre-processed to handle missing values and incorrect entries. Missing values were imputed, and inaccurate entries were removed. The pre-processed dataset was split into 80% training and 20% test sets to ensure model generalizability.

2.3. Model Fitting

Six machine learning techniques were employed for HF survival prediction: Random Forest (KF) Classifier, Multi-Layer Perceptron Network, Support Vector Machine, K-Nearest Neighbors, Naive Bayes, and Logistic Regression. These classifiers were chosen based on their performance across the entire dataset.

2.4. Class Imbalance

To address the class imbalance in the dataset (203 survived, 96 died), the Synthetic Minority Over-sampling Technique (SMOTE) was used to oversample the minority class, ensuring a balanced dataset during model training (See Figures 1 and 2) [24].

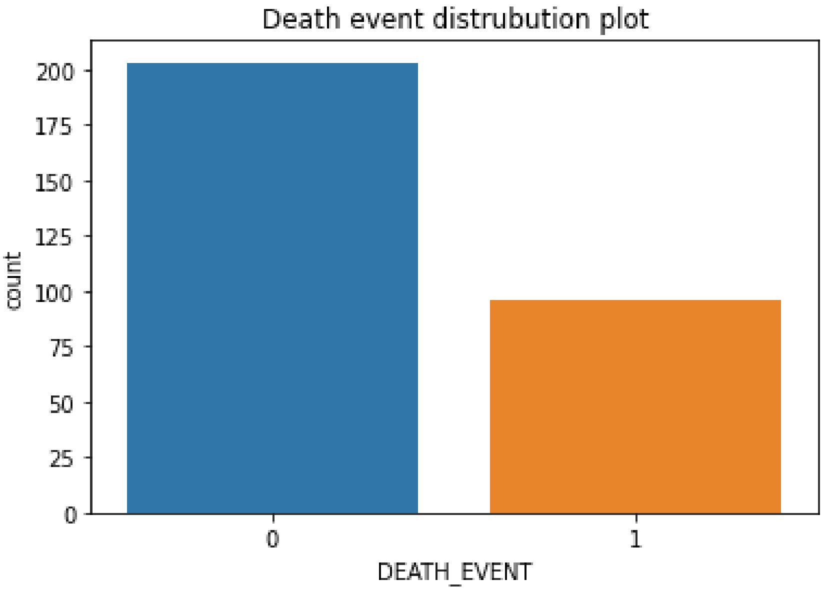


Figure 1. Class Distribution of Death Event on Original Dataset

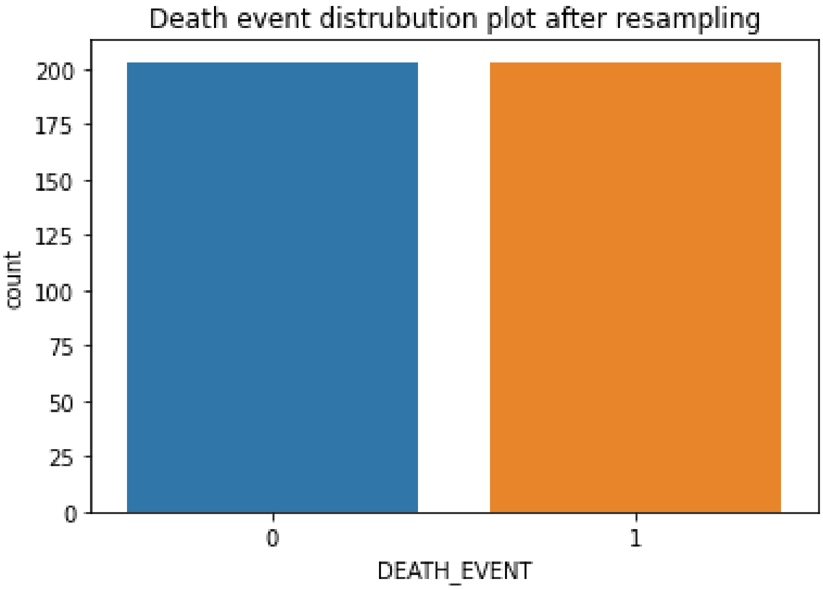


Figure 2. Class Distribution after Applying SMOTE to fix Imbalance

2.5. Feature Engineering

Feature ranking was performed using conventional univariate biostatistics analysis (Mann-Whitney U and Chi-Square tests) and machine learning analysis. The top-ranking features were identified for model training.

2.6. Performance Metrics

Performance evaluation was done using various metrics, considering the class imbalance in the dataset. Metrics included accuracy, sensitivity (Recall), Matthews's correlation coefficient (MCC), and receiver operating characteristics-area under the curve (ROC-AUC). These metrics provided a comprehensive assessment of the classifiers' ability to predict HF survival accurately.

Figure 3, shows the flowchart of the research methodology through which the objectives of this study were achieved.

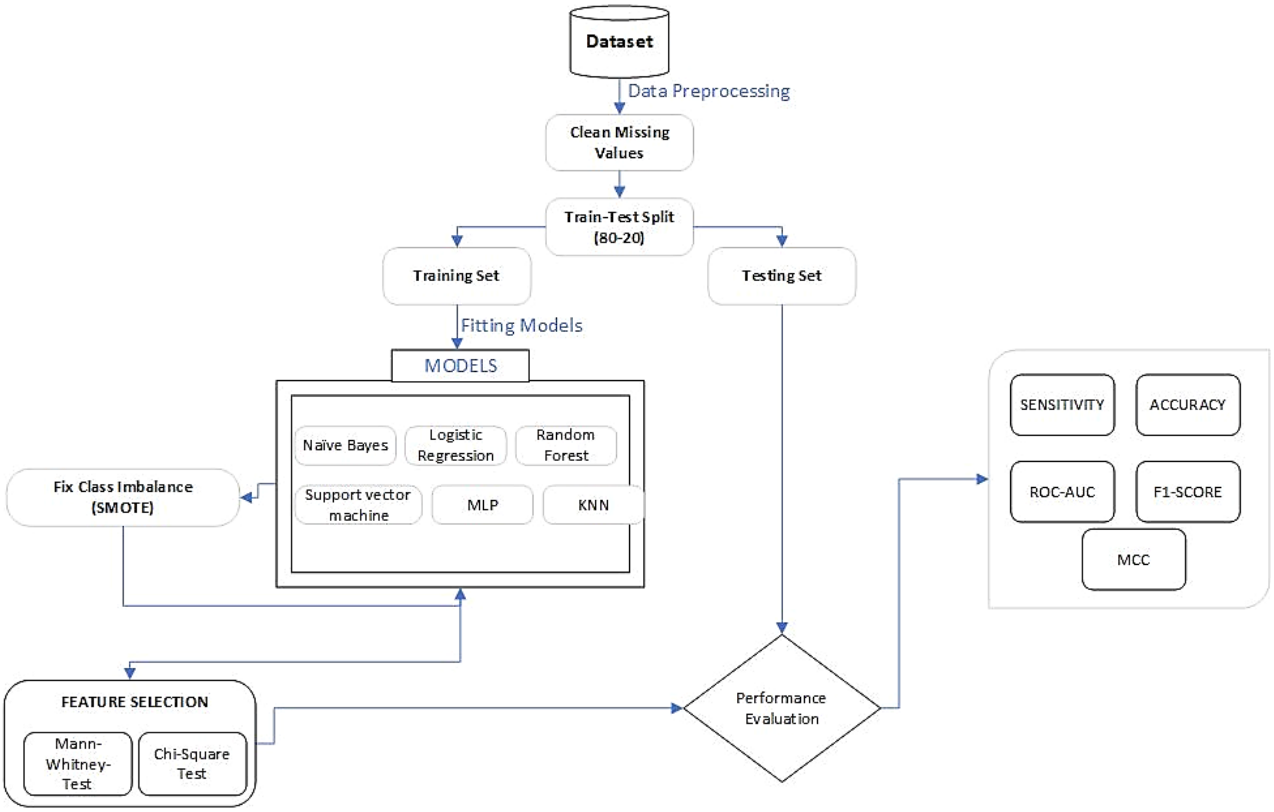


Figure 3. Flowchart of Research Methodology

3. Results and Discussion

3.1. Exploring Key Features of the Dataset through Informative Visualization Plots

The visualization plots provided valuable insights into the dataset. The violin plot (Figure 4) displayed the distribution of ejection fraction for different groups based on age and high blood pressure. It indicated that older individuals with high blood pressure tended to have lower ejection fractions. The cat plot (Figure 5) showed a relationship between death events and ejection fraction, revealing that individuals who died from heart failure generally had lower ejection fractions than survivors. The scatter plot (Figure 6) demonstrated a negative correlation between platelet levels and age.

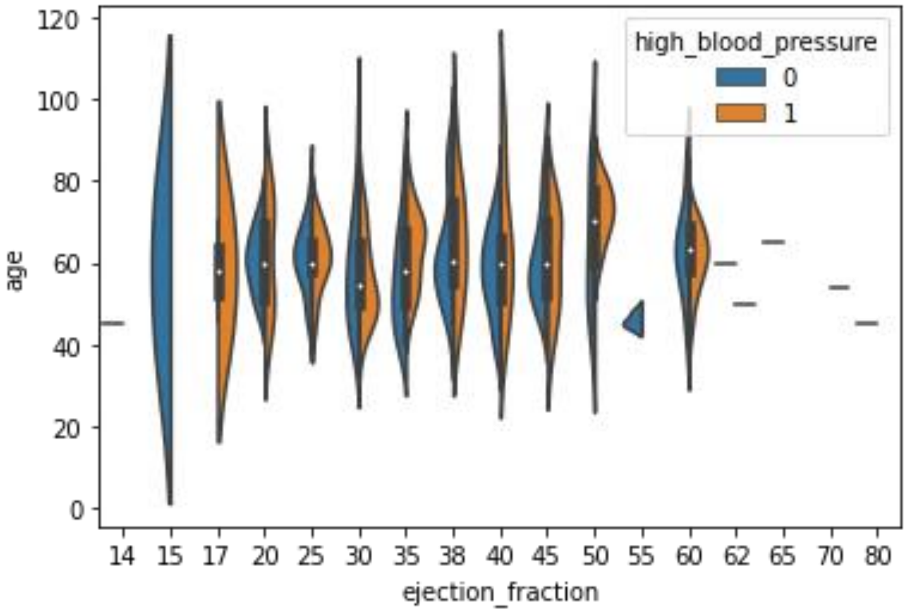


Figure 4. Ejection fraction for different groups

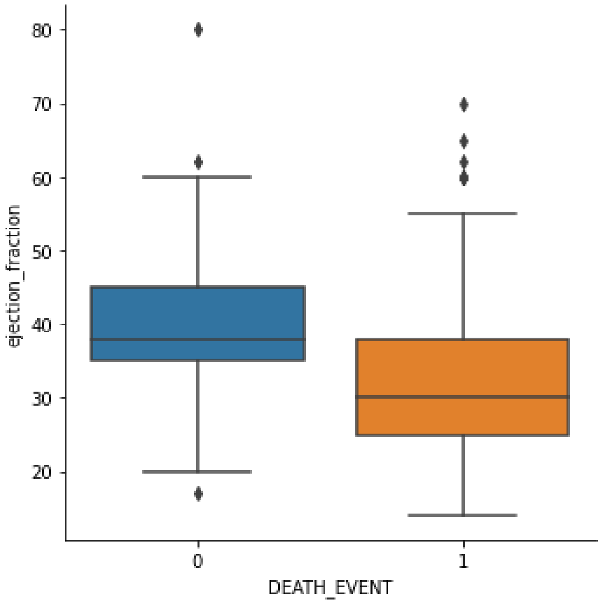


Figure 5. The relationship between death events and ejection fraction

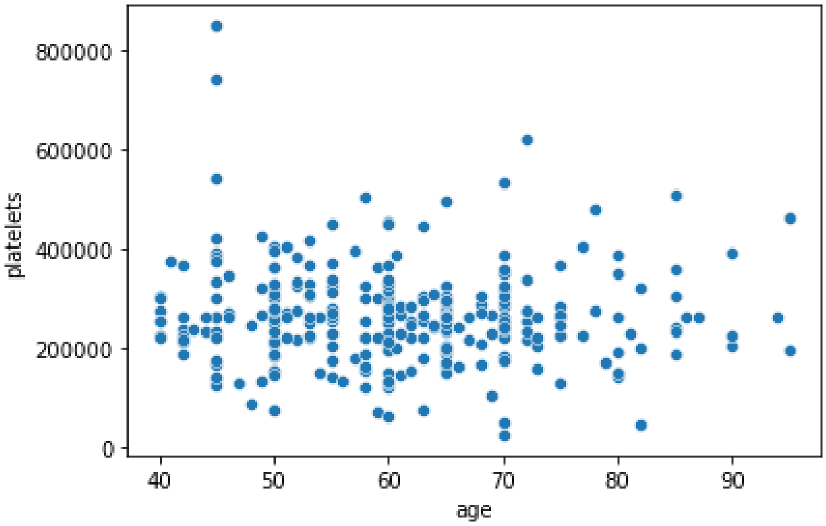


Figure 6. Relationship between platelet levels and age

3.2. Performance Comparison of Models Developed Using Standard RF Classifier, SVM, KNN, LR, Naïve Bayes, and MLP

The performance measures of the models in Table 1 demonstrated varying levels of accuracy, sensitivity, specificity, and area under the ROC curve (ROC-AUC). It also showed the disparity between the sensitivity and specificity measurements which brings to the understanding that the prediction is skewed in favor of the majority class in the dataset. Moreover, the classification algorithms are not built to handle such imbalanced situations. Meanwhile, the Random Forest (RF) classifier exhibited the highest accuracy (73.92%) and ROC-AUC (67.39%) among the models tested. However, all models showed limitations in handling class imbalance, leading to skewed predictions.

Table 1. Performance measures (%) obtained using standard classification algorithms

Evaluation Metrics	SVM	KNN	LR	NB	RF	MLP
Sensitivity	58.0	53.0	80.00	58	73.00	58
Specificity	97.05	91.65	93.1	82.24	85.73	82.12
G-mean	26.44	35.72	56.12	64.25	64.31	60.23
Accuracy	58.00	53.56	80.23	73.23	73.92	58.00
MCC	11.57	9.43	35.16	34.59	37.84	40.23
ROC-AUC	53.23	53.11	63.71	66.62	67.39	67.30

3.3. Performance Comparison of Models Developed Using Standard RF Classifier, SVM, KNN, LR, Naïve Bayes, and MLP after Applying SMOTE

Table 2 revealed that after applying SMOTE to address class imbalance, the models performance improved significantly. While the RF classifier achieved the highest accuracy (92.93%) and ROC-AUC (89.43%). Sensitivity and specificity also improved for most models, indicating better predictive capabilities.

Table 2. Performance measures (%) obtained after applying SMOTE

Empty Cell	SVM	KNN	LR	NB	RF	MLP
Sensitivity	56.42	33.47	72.00	83.00	89.00	57.00
Specificity	69.5	73.46	71.98	73.93	79.34	70.00
G-mean	61.98	48.24	68.61	65.32	68.41	62.95
Accuracy	60.21	57.54	73.00	83.91	92.93	55.78
MCC	25.00	7.25	35.97	32.01	89.68	34.01
ROC-AUC	62.96	53.47	68.8	66.17	89.43	54.00

3.4. Feature Ranking

Tables 3 and 4 (biostatistics tests) showed the features of the dataset that they are both significant for the serum creatinine and ejection fraction variables. Comparing the significant results of the biostatistics feature ranking test with the machine learning ranking test in Figure 5, it can be seen that the variables serum creatinine and ejection fraction also exists among the trio as the most significant features of our modeling. This is related to the study by Chicco & Jurman [7]. With this, we conclude that the features serum creatinine and ejection fraction are the most significant models for predicting HF survival in the observed patients.

Table 3. Features from the Mann-Whitney test

Features	p-value
Time	6.623261e-16
Serum Creatinine	1.820929e-10
Ejection Fraction	3.454241e-06
Serum Sodium	4.657701e-03
Age	5.859484e-03

Table 4. Features from the Chi-square test

Features	p-value
Ejection Fraction	0.000092
Serum Creatinine	0.000692
Time	0.000956

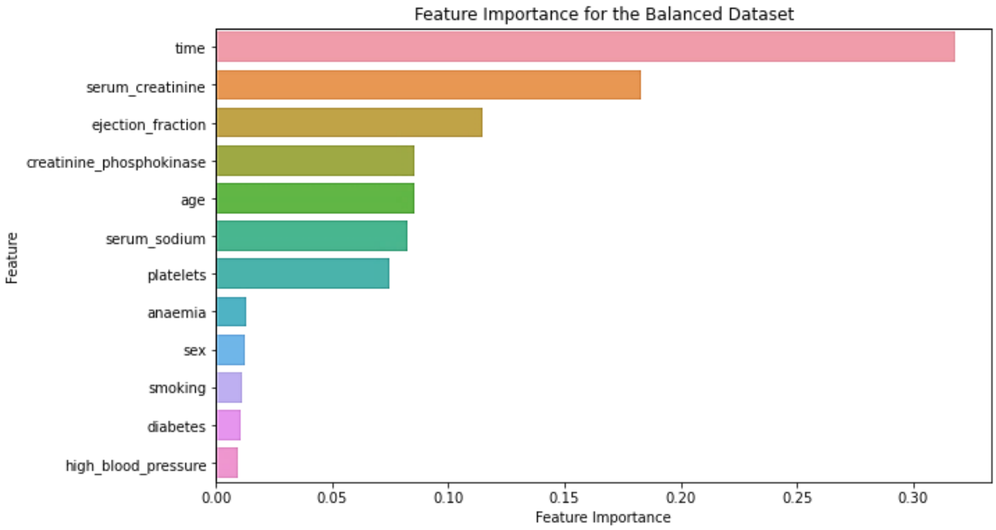
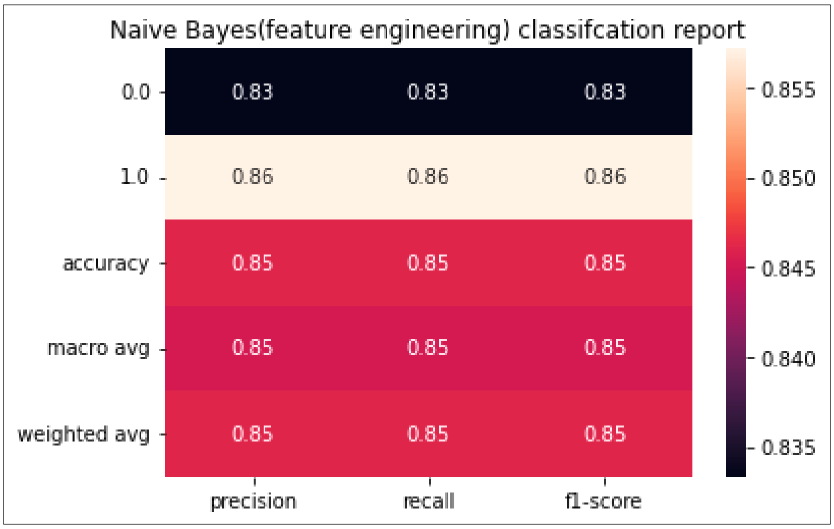
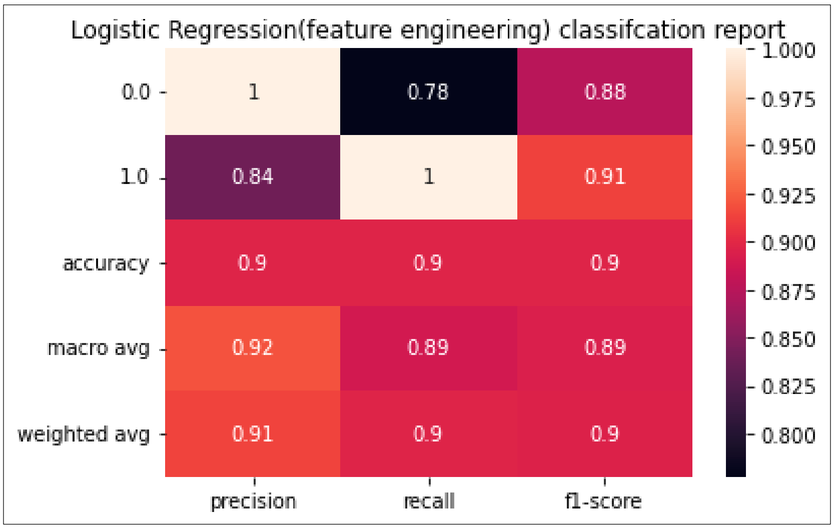


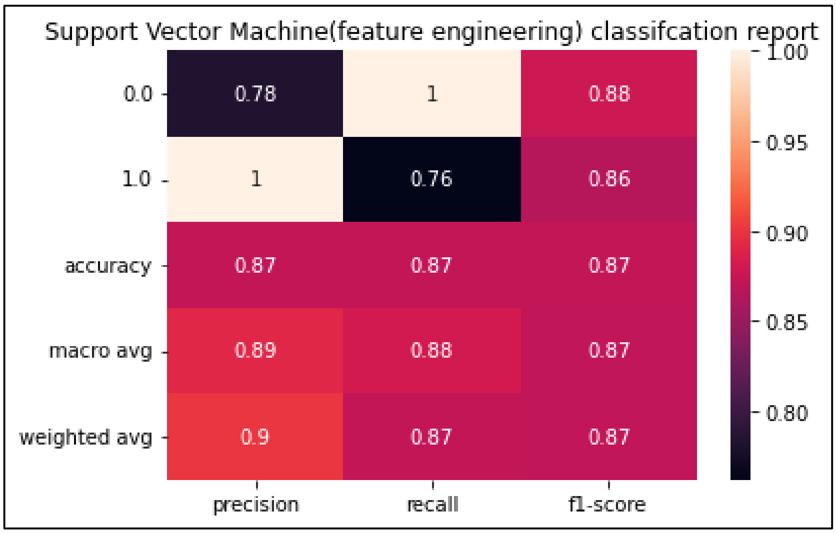
Figure 5. Feature Importance Graph from Random Forest Classifier

3.5. Remodeling of Models with Selected Features

Retraining the models with the selected features (serum creatinine and ejection fraction) significantly improved their accuracy and predictive capabilities as it reduced the dimensionality of the data. Looking at the classification reports in Figure 6, there is a massive increase in precision, sensitivity, and accuracy across all models, but the RF classifier remains the top-performing model with accuracy above 90%. This is in agreement with the results by Zhang et al (2020) that RF classifier has a high accuracy in the risk predictions in patients with cardiovascular disease.







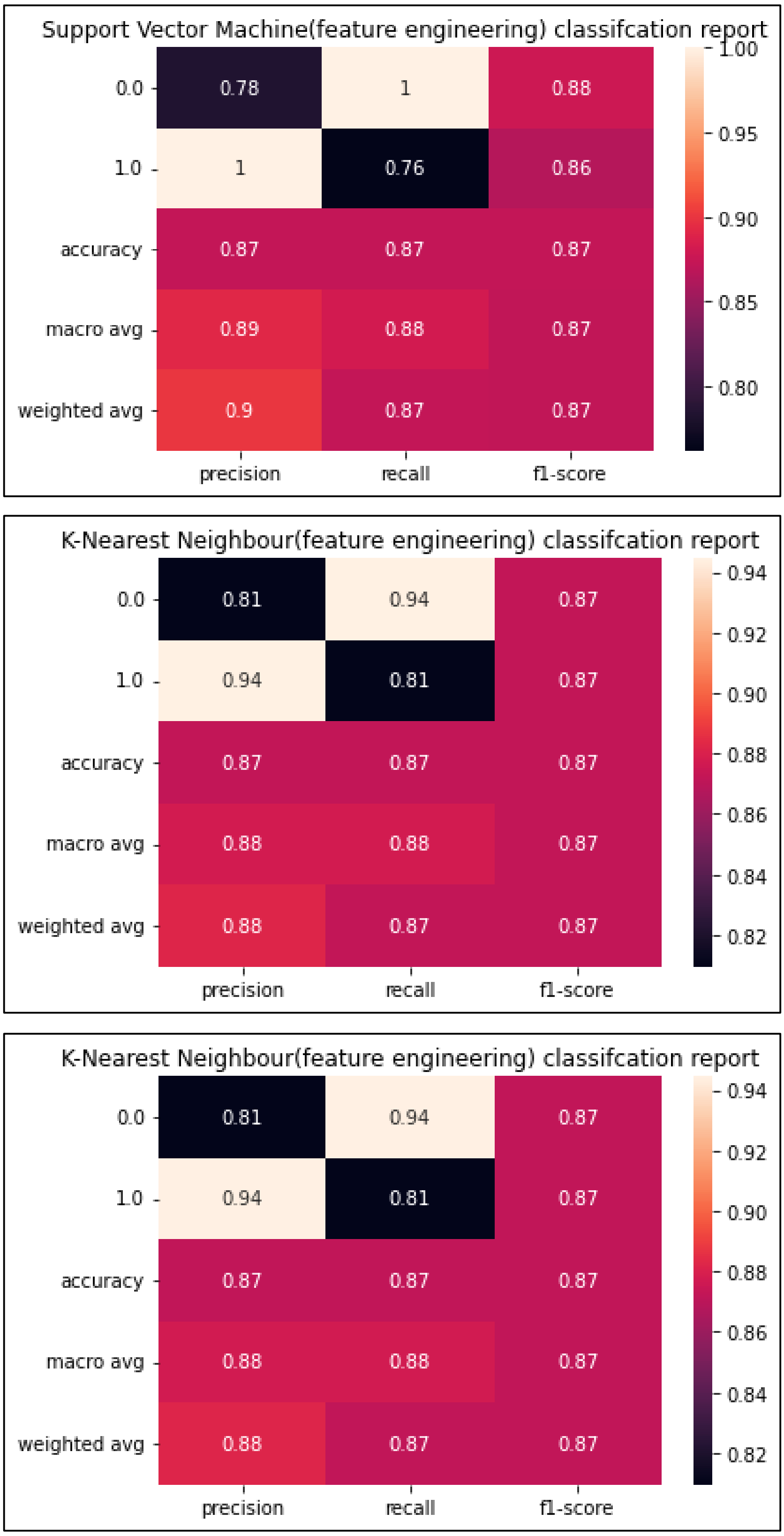


Figure 6. Classification Report of Models Retrained with Significant Features

4. Conclusion

This study utilized the medical data of 299 heart failure patients from the Faisalabad Institute of Cardiology and the Allied Hospital in Faisalabad, Pakistan. The following six machine learning models were trained and evaluated: Random Forest (KF) Classifier, Multi-Layer Perceptron Network, Support Vector Machine, K-Nearest Neighbors, Naive Bayes, and Logistic Regression. However, the initial results indicated that the models were biased toward predicting most patients who did not die from heart failure. By addressing the class imbalance in the data with the synthetic minority oversampling technique (SMOTE), the performance of the six models improved significantly. Among the six models tested, the Random Forest classifier and Naïve Bayes showed the best accuracy in predicting heart failure survival. Furthermore, the study identified the most significant features for survival prediction through the Mann-Whitney and Chi-Square tests: Serum Creatinine and Ejection Fraction. Restructuring the six models using only these two features resulted in accuracy levels above 90% for the Random Forest Classifier, making it the most effective model for future predictions. This is in line with previous research by Chicco & Jurman (2020) [7], Patidar et al. (2020) [25], and Li et al. (2022) [26]. Therefore, the study demonstrated that machine learning methods could effectively predict the survival of patients with heart failure through the binary classification of electronic health records of patients with cardiovascular heart diseases. That is, the Random Forest Classifier would be helpful in improving prediction and treatment, leading to better patient outcomes and mortality rates associated with heart failure.

5. Declarations

5.1. Author Contributions

Conceptualization, O.E.O. and O.B.A.; methodology, O.E.O.; software, O.E.O.; validation, O.E.O., O.B.A., and O.K.B.; formal analysis, O.E.O.; investigation, O.E.O. and O.B.A.; resources, O.K.B.; data curation, O.E.O.; writing—original draft preparation, O.E.O. and O.B.A.; writing—review and editing, O.K.B.; visualization, O.E.O.; supervision, O.B.A.; project administration, O.B.A. All authors have read and agreed to the published version of the manuscript.

5.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

5.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

5.4. Acknowledgements

We appreciate UCI Machine Learning Repository for making Heart Failure Clinical Records Data Set available to us.

5.5. Institutional Review Board Statement

Not applicable.

5.6. Informed Consent Statement

Not applicable.

5.7. Declaration of Competing Interest

The authors declare that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

6. References

[1] CDC. (2023). Heart Failure. Centers for Disease Control and Prevention. Centers for Disease Control and Prevention. U.S. Department of Health and Human Services. Available online: https://www.cdc.gov/heartdisease/heart_failure.htm (accessed on March 2023).

[2] WebMD. (2023). Heart Attack. WebMD LLC. Available online: https://www.webmd.com/heart-disease/heart-disease-heart-attacks (accessed on March 2023).

[3] CDC. (2023). Health Topics – Heart Disease and Heart Attack. Centers for Disease Control and Prevention. U.S. Department of Health and Human Services. Available online: https://www.cdc.gov/policy/polaris/healthtopics/heartdisease/index.html (accessed on March 2023).

[4] WHO. (2023). Cardiovascular diseases (CVDs). World Health Organization. Available online: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed on March 2023).

[5] Sawano, M., Shiraishi, Y., Kohsaka, S., Nagai, T., Goda, A., Mizuno, A., Sujino, Y., Nagatomo, Y., Kohno, T., Anzai, T., Fukuda, K., & Yoshikawa, T. (2018). Performance of the MAGGIC heart failure risk score and its modification with the addition of discharge natriuretic peptides. ESC Heart Failure, 5(4), 610–619. doi:10.1002/ehf2.12278.

[6] Cooper, L. B., & Hernandez, A. F. (2015). Assessing the Quality and Comparative Effectiveness of Team-Based Care for Heart Failure. Who, What, Where, When, and How. Heart Failure Clinics, 11(3), 499–506. doi:10.1016/j.hfc.2015.03.011.

[7] Chicco, D., & Jurman, G. (2020). Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC Medical Informatics and Decision Making, 20(1), 16. doi:10.1186/s12911-020-1023-5.

[8] Rahimi, K., Bennett, D., Conrad, N., Williams, T. M., Basu, J., Dwight, J., Woodward, M., Patel, A., McMurray, J., & MacMahon, S. (2014). Risk Prediction in Patients with Heart Failure. JACC: Heart Failure, 2(5), 440–446. doi:10.1016/j.jchf.2014.04.008.

[9] Edgar, T.W., & Manz, D.O. (2017). Research methods for cyber security. Research Methods for Cyber Security: Syngress, 428.

[10] Soori, M., Arezoo, B., & Dastres, R. (2023). Artificial intelligence, machine learning and deep learning in advanced robotics, a review. Cognitive Robotics, 3, 54–70. doi:10.1016/j.cogr.2023.04.001.

[11] Pocock, S. J., Ariti, C. A., McMurray, J. J. V., Maggioni, A., Køber, L., Squire, I. B., Swedberg, K., Dobson, J., Poppe, K. K., Whalley, G. A., & Doughty, R. N. (2013). Predicting survival in heart failure: A risk score based on 39 372 patients from 30 studies. European Heart Journal, 34(19), 1404–1413. doi:10.1093/eurheartj/ehs337.

[12] Tohyama, T., Ide, T., Ikeda, M., Kaku, H., Enzan, N., Matsushima, S., Funakoshi, K., Kishimoto, J., Todaka, K., & Tsutsui, H. (2021). Machine learning-based model for predicting 1 year mortality of hospitalized patients with heart failure. ESC Heart Failure, 8(5), 4077–4085. doi:10.1002/ehf2.13556.

[13] Newaz, A., Ahmed, N., & Shahriyar Haq, F. (2021). Survival prediction of heart failure patients using machine learning techniques. Informatics in Medicine Unlocked, 26, 100772. doi:10.1016/j.imu.2021.100772.

[14] Zahid, F. M., Ramzan, S., Faisal, S., & Hussain, I. (2019). Gender based survival prediction models for heart failure patients: A case study in Pakistan. PLoS ONE, 14(2), 1-10. doi:10.1371/journal.pone.0210602.

[15] Mamun, M., Farjana, A., Al Mamun, M., Ahammed, M. S., & Rahman, M. M. (2022). Heart failure survival prediction using machine learning algorithm: am I safe from heart failure?. In 2022 IEEE World AI IoT Congress (AIIoT), 194-200. doi:10.1109/AIIoT54504.2022.9817303

[16] Handelman, G. S., Kok, H. K., Chandra, R. V., Razavi, A. H., Lee, M. J., & Asadi, H. (2018). eDoctor: machine learning and the future of medicine. Journal of Internal Medicine, 284(6), 603–619. doi:10.1111/joim.12822.

[17] Al Mehedi Hasan, M., Shin, J., Das, U., & Yakin Srizon, A. (2021). Identifying Prognostic Features for Predicting Heart Failure by Using Machine Learning Algorithm. ACM International Conference Proceeding Series, 40–46. doi:10.1145/3460238.3460245.

[18] Al'Aref, S. J., Anchouche, K., Singh, G., Slomka, P. J., Kolli, K. K., Kumar, A., Pandey, M., Maliakal, G., Van Rosendael, A. R., Beecy, A. N., Berman, D. S., Leipsic, J., Nieman, K., Andreini, D., Pontone, G., Schoepf, U. J., Shaw, L. J., Chang, H. J., Narula, J., … Min, J. K. (2019). Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. European Heart Journal, 40(24), 1975–1986. doi:10.1093/eurheartj/ehy404.

[19] Benjamins, J. W., Hendriks, T., Knuuti, J., Juarez-Orozco, L. E., & van der Harst, P. (2019). A primer in artificial intelligence in cardiovascular medicine. Netherlands Heart Journal, 27(9), 392–402. doi:10.1007/s12471-019-1286-6.

[20] Fernandes, K., Chicco, D., Cardoso, J. S., & Fernandes, J. (2018). Supervised deep learning embeddings for the prediction of cervical cancer diagnosis. PeerJ Computer Science, 2018(5), 1-20. doi:10.7717/peerj-cs.154.

[21] Virani, S. S., Alonso, A., Aparicio, H. J., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., Carson, A. P., Chamberlain, A. M., Cheng, S., Delling, F. N., Elkind, M. S. V., Evenson, K. R., Ferguson, J. F., Gupta, D. K., Khan, S. S., Kissela, B. M., Knutson, K. L., Lee, C. D., Lewis, T. T., … Tsao, C. W. (2021). Heart Disease and Stroke Statistics - 2021 Update: A Report from the American Heart Association. Circulation, 143(8), E254–E743. doi:10.1161/CIR.0000000000000950.

[22] Kaggle. (2019). Heart Failure Prediction. Kaggle. Available online: https://www.kaggle.com/datasets/andrewmvd/heart-failure-clinical-data (accessed on March 2023).

[23] UCI Machine Learning Repository (2020). Heart Failure Clinical Records. Irvine, California, United States. Available online: https://archive.ics.uci.edu/ml/datasets/Heart+failure+clinical+records (accessed on March 2023).

[24] Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). Learning from imbalanced data sets Springer, Vol. 10, 377. doi:10.1007/978-3-319-98074-4.

[25] Patidar, S., Kumar, D., & Rukwal, D. (2022). Comparative Analysis of Machine Learning Algorithms for Heart Disease Prediction. Advances in Transdisciplinary Engineering, 27, 64–69. doi:10.3233/ATDE220723.

[26] 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. doi:10.1109/ACCESS.2020.3001149.