OkIP Tiako P.F. (ed) *Competitive Tools, Techniques, and Methods*. Chronicle of Computing. OkIP. AHIT24#15

https://doi.org/10.55432/978-1-6692-0007-9_4

Heart Failure Prediction using Machine learning with Metaheuristic feature selection techniques.

Omar Elgendy[a], Ali Bou Nassif[a], Bassel Soudan[a]

University of Sharjah, Sharjah, United Arab Emirates

ABSTRACT

Heart failure is a major complication of cardiovascular diseases (CVDs), which are the leading cause of mortality globally. However, early treatment and detection of heart failure may increase the chance of survival. Using current clinical data, machine learning (ML) algorithms present an effective approach for predicting heart failure. Utilizing ML algorithms and feature selection using metaheuristic methods, we present a novel framework in this research for the prediction of heart failure. We employ several ML algorithms and perform feature selection using four meta-heuristic techniques: Grey Wolf Optimization (GWO), Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Ant Colony Optimization (ACO). The performance of each combination is evaluated and compared based on the F-score metric, which we aim to maximize. (PSO) has achieved the highest when choosing the relevant features which increased the overall accuracy from 0.83 to 0.90. The results indicate that our proposed framework can effectively identify relevant features and improve the predictive performance of ML algorithms for heart failure. Furthermore, we provide a comprehensive comparison of the meta-heuristic techniques, highlighting their advantages and limitations for feature selection in heart failure prediction.

Keywords: Machine Learning, Genetic Algorithm, Feature Selection, Ant Colony Optimization, Particle Swarm Optimization

I. INTRODUCTION

With millions of individuals impacted each year, heart disease continues to be one of the major causes of death and morbidity globally. Strategies for early identification and prevention are crucial for lowering the burden of heart disease and improving patient outcomes and accuracy of detection plays an important role in early diagnosis and treatment. The annual Statistical Update on heart diseases reveals crucial data on cardiovascular health (CVH) in the United States. Between 2013 and March 2020, the average overall CVH score for US children (16-19 years) was 73.6, while adults scored 65.2. Ideal CVH is associated with life expectancy gains of 5.50 years for males and 4.20 years for females with cardiometabolic diseases. As of July 1, 2022, cumulative COVID-19 deaths reached 1,014,620 in the US, impacting life expectancy. Disparities in tobacco use persist among various demographics. Meanwhile, physical activity, nutrition, and obesity rates influence CVH, with 71.2% of adults being overweight or obese. High blood pressure affects 46.7% of US adults, while cholesterol levels remain concerning [1]. Heart disease is a complex medical issue with many risk factors and symptoms, making early detection and prevention more difficult. Technology has advanced quickly, and data-driven methods have made it possible for more efficient diagnosis and treatment methods like machine learning and artificial intelligence (AI). By utilizing the power of data and offering insights into patterns and linkages that may be challenging to discover using conventional techniques, these strategies have the potential to alter how physicians and researchers identify, understand, and manage heart disease [2].

To find patterns and generate predictions with astounding accuracy, machine learning algorithms may examine complex datasets containing a variety of factors. Based on a variety of variables, including age, sex, lifestyle choices, and biomarkers, these algorithms can be used to estimate a patient's risk of acquiring a cardiovascular illness in the setting of heart disease. Additionally, AI-driven diagnostic technologies can help doctors make better judgments, which will eventually improve patient care and ease the strain on healthcare systems [3].

In this research paper, we aim to develop an approach to predict heart failure using structured data and apply feature selection techniques using meta-heuristic optimization algorithms. By optimizing the choice of features and maximizing the f-score, our objective is to create accurate and efficient predictive models that can revolutionize cardiac care and save countless lives.

II. RELATED WORK

A. Machine Learning

Machine learning is a rapidly developing technical area that is utilized in many sectors to help people make decisions based on the best available data. To strike a balance between privacy and justice while promoting new approaches, policymakers and the general public's talks regarding data and machine learning are essential. Machine learning is now a widely utilized and useful technology with a wide range of applications, from social science to computer vision. Instead of manual programming, it uses examples to instruct systems. Industries involved with data-intensive challenges are also impacted [4]. Numerous fields use machine learning algorithms and their applications. Unsupervised learning, which includes techniques like clustering and dimension reduction, is a subfield of machine learning that focuses on classification and result prediction. Machine learning has recently made advancements that address issues including distributed processing, communication limitations, and computing design.

Presented at the 2024 OkIP International Conference on Advances in Health Information Technology (AHIT) in Oklahoma City, OK, USA, and Online, on October 3, 2024

Researchers are striving to create algorithms that operate well in these settings, taking resource trade-offs into account and protecting privacy using techniques like differential privacy. Machine learning enables pattern recognition and predictions in massive datasets and has applications in fields including statistics, computer science, and big data [5].

B. *Genetic Algorithm:*

Before The genetic algorithm imitates natural selection for problem-solving and optimization because it was influenced by Charles Darwin's theory of natural evolution. It seeks to give the optimal solution for diverse optimization problems, including those with discontinuous, nondifferentiable, stochastic, or highly nonlinear objective functions, by gradually developing a population of alternative solutions. Creating an initial population, assessing fitness, choosing parents, having offspring, evaluating the new generation, and repeating the process until a stopping requirement is reached are the six basic processes that make up the algorithm's operation. As fitter people pass on their traits to younger generations, the population gradually evolves toward an ideal solution [6]. Due to their advantages, such as not requiring derivative knowledge, being quicker and more efficient than traditional techniques, and providing multiple viable solutions, genetic algorithms have grown in popularity[1]. Due to their superior parallel capabilities, they are particularly helpful for optimizing continuous and discrete functions as well as multi-objective problems. Genetic algorithms do have limitations, though. They can be computationally expensive to calculate fitness values and may not be appropriate for simple problems with readily available derivative information. Due to its stochastic nature, the solution's effectiveness or optimality cannot be guaranteed, and erroneous application may yield less-than ideal results.

C. *Particle Swarm Optimization:*

Kennedy and Eberhart invented the optimization technique known as particle swarm optimization (PSO) in 1995. They were motivated by the movement of social groupings like flocks of birds or schools of fish [7]. Without relying on the gradient or other differential forms of the objective function, it finds the optimum solution within a problem space. An initial population of particles with random places and velocities is created by the algorithm, which then assesses their costs and updates the population's positions depending on the best particle and their current velocities. Until a termination criterion is satisfied, this process is repeated. Due to its simplicity, quick convergence, noise resistance, and parallelizability, PSO is well-liked. It has drawbacks, though, include being sensitive to parameter choice, being stuck in local optima, and maybe needing numerous iterations to produce satisfying results. PSO works well for dynamic and complex optimization problems, but the parameters used can have a big impact on how well it works [8].

D. *Ant Colony Optimization:*

An optimization algorithm called Ant Colony Optimization (ACO) is based on the foraging strategies of ants. It mimics an artificial ant colony that explores the search space repeatedly while assessing potential solutions and updating pheromone trails to gauge how desirable they are [9]. The approach discovers global optimum, converges quickly, is robust, can handle non-linear and non-convex problems, and is simple to parallelize. It can tackle complex optimization issues with numerous variables and constraints and adapt to changing problem environments. ACO has certain drawbacks, though, such as the need for a considerable amount of memory to keep pheromone trails and the computational cost of solving problems with huge search spaces and sensitivity to parameter choices. Despite these difficulties, ACO continues to be a well-liked and successful optimization algorithm for many practical applications [10].

E. *Artificial Bee Colony:*

In 2005, Karaboga created the Artificial Bee Colony (ABC) algorithm, an optimization technique based on the foraging habits of honeybees [11]. Workers, observers, and scout bees make up the algorithm, which can solve maximization and minimization issues. Initialization, employed bees, observer bees, scout bees, and termination phases make up the process, which produces the optimal solution discovered throughout execution. ABC has several benefits, including its ease of use, adaptability, and ability to handle various optimization issues without the need for additional problem-specific parameters. It also has drawbacks, such as the tendency to become caught in local optima because it relies on local search techniques and the lack of a convergence guarantee, which may necessitate extra work to decide when to stop the algorithm. ABC is nevertheless a useful optimization technique for a variety of applications despite these drawbacks [12].

Authors in [13] propose a novel wrapper-based feature selection approach for predicting heart failure, enhancing the Multilayer Perceptron architecture using an Adaptive Particle Swarm Grey Wolf Optimization (APSGWO). The research uses heart failure patient data from the Faisalabad Institute of Cardiology and Allied Hospital, Pakistan. This data was preprocessed using the Interquartile Range (IQR) method to remove outliers. The model's performance was evaluated using accuracy, precision, recall, and F1-score metrics. The results demonstrated an improved prediction accuracy with fewer features, indicating a more efficient and effective method for early heart failure diagnosis. Particularly, the machine learning algorithms logistic regression (LR) and support vector machine (SVM) showed an accuracy of 80%. The study demonstrates the potential of machine learning algorithms in medical diagnostics when paired with innovative feature selection methods.

Authors in [14] The study investigates the viability of employing artificial intelligence (AI) algorithms to analyze vital sign data from electronic health records (EHR) to forecast in-hospital cardiac arrest (CA). Using information from the MIMIC III database, the study assessed six AI algorithms: CNN, KNN, MLP, NB, RF, SVM, and NB. With just one hour of data prior to the start of CA, the Random Forest (RF) algorithm routinely produced the best predictions, surpassing 80%. On the other hand, Naive Bayes had the worst results. There was not much of a performance increase when the observation window was extended to 12 hours. The study shows that utilizing conventional EHR data, AI can accurately identify imminent CA within the next 60 minutes, allowing for prompt medical intervention. However, generalizability to ordinary ward settings is limited due to the study's dependence on high-frequency ICU data, and its concentration on short-term prediction limits its wider usefulness. The study's contribution lies in identifying the optimal algorithm (RF) and the minimum data required (1 hour) for accurate CA prediction, providing a practical approach for early intervention in clinical settings. Limitations include dependency on high-frequency ICU data and the restriction to short-term predictions, which may not be as applicable in less monitored environments

The authors in [15] describe a Faster Region-based Convolutional Neural Network (FRCNN) that has been Honey Badger Algorithm (HBA) improved for effective CHF prediction. The method focuses on extracting QRS complexes from noisy ECG data using the Discrete Cosine Transform (DCT) and fast Fourier transformations (FFT), as well as preprocessing the signals using the Delayed Normalized Least Mean Square (DNLMS) method. The model showed exceptional accuracy scores of 98.65%, 97.81%, 98.5%, and 98.2% when tested with two datasets. The HBA, a technique for global optimization based on honey badger feeding behavior, was essential in promoting the study and utilization of potential solutions. Additionally, in order to lower the Mean Squared Error (MSE) value, the researchers used adaptive filtering for ECG signal processing utilizing the LMS and NLMS algorithms. The results highlight the value of machine learning in cardiac diagnosis, particularly in spotting potentially fatal sequences in intensive care unit patients.

In [16] authors suggest an improved Random Forests classifier with feature reduction in their study to forecast heart failure from gene expression data. 46 healthy controls and 111 patients who had ST-elevation myocardial infarction (STEMI) were the subjects of the research. To eliminate batch effects, the data was preprocessed using the ComBat approach. Feature ranking and feature deletion were then used to verify the model. To create a balanced dataset from unbalanced data, the authors used undersampling, over-sampling, and Random Forests. An MS technique was used to choose the model hyper-parameters. The receiver operating characteristic area under the curve for the model was 0.918, and the Matthews Correlation Coefficient was +0.87. A literature study and a Gene Set Enrichment Analysis were used to confirm the results, which also revealed the top 5 protein-coding genes linked to heart failure.

In [17], the authors develop a hybrid decision support system to identify myocardial infarction (MI) and constantly track blood pressure with the goal of delivering accurate management and prompt heart disease detection. To account for linear and non-linear correlations between variables, the model uses a combination of control charts and prediction models, such as artificial neural networks (ANNs). Data dynamics and uncertainty are also lessened by the hybrid model. The model was verified using an actual medical case study from Iran, which showed significant accuracy in identifying the severity of MI. The model's ability to predict blood pressure and the severity of a MI was further examined using MATLAB 2016b, and the results showed an accuracy rate of 81.25%. In terms of actual systolic and diastolic blood pressure readings, the system's overall accuracy was determined to be 100%. The approach developed by the authors presents bright potential for healthcare analytics in heart disease.

The study in [18] presents a hybrid federated learning technique based on classifiers for predicting cardiovascular disease (CVD). The framework combines a support vector machine with modified artificial bee colony optimization to increase accuracy and lower classification error. Healthcare systems' ingrained privacy concerns are addressed by the federated matched averaging (FedMA) method. The authors show that their model outperforms currently available federated learning algorithms like FedAvg-SVM and FedMA-SVM by achieving 93.8% accuracy within 4500 communication cycles. According to the study's findings, the suggested approach successfully improves prediction accuracy while also ensuring data privacy and minimizing the number of communication rounds required, hence reducing computing costs and communication overheads. The discovery presents exciting opportunities for privacy-aware cardiac disease prediction and IoMT based healthcare solutions.

In [19] authors use SVM, Naive Bayes, Logistic Regression, Decision Trees, and KNN on a dataset from the UCI repository to provide a machine learning-based method for anticipating cardiac failure. Risk factors that may be changed include those like smoking, diabetes, obesity, and hypertension. SVM was shown to be very successful in producing hyperplanes and classifying data. The system outperformed previous approaches with an accuracy of 86.42% using SVM. 20% of Covid-19 patients have cardiac damage, according to the study, which emphasizes the need of early diagnosis and treatment. The predictive system was assessed using the criteria of accuracy, precision, recall, and F-1 score, demonstrating its dependability and potential to enhance patient outcomes and healthcare efficiency.

The study in [20] suggests an integrated decision support system that predicts the risk of heart failure with a classification accuracy of 95.55% by combining machine learning models with a grid of stacked autoencoders. This performs better than current feature transformation techniques. Each autoencoder in the grid has a unique architecture and code size, and they all work together to extract additional features from the original input. By maximizing model capacity and employing quality characteristics, the study solves underfitting, overfitting, and overfitting to testing data concerns. The integrated approach performs better when paired with machine learning models than PCA, KPCA, and SPCA feature transformation techniques, according to experiments. Additionally, the study stresses the need of avoiding sacrificing training accuracy to increase testing accuracy and use the Mean Percentage Error (MPE) to assess model generalization.

III. Methodology

Our goal is to predict the heart failure in the best optimized and efficient way. To predict heart failure, we use feature selection approaches using an open-source structured dataset in the methodology portion of this study report. By using a method that maximizes the f-score to optimize the selection of features, we can make sure that the resultant prediction models are both precise and effective. The quality of cardiac treatment is expected to increase. In this section, we will talk about the details of the dataset, models, and methodology of applying the feature selection using meta heuristic algorithms.

A. Dataset:

In this study, we utilize the Heart Failure Prediction Dataset 2, an open-source dataset available on Kaggle, to predict heart failure. The dataset contains 918 observations and consists of the following attributes: Age, Sex, chest pain type, resting blood pressure, Cholesterol, fasting blood sugar, resting ECG results, maximum heart rate achieved, exercise-induced angina, old peak, the slope of the peak exercise ST segment, and heart disease as the output class as a binary classification for 1 if there is a heart disease and 0 if there is no heart diseases. This comprehensive dataset was created by combining five independent heart disease datasets, namely Cleveland, Hungarian, Switzerland, Long Beach VA, and Stalog (Heart) Data Set, resulting in the largest dataset available for heart disease research. Although it has small number of observations, this dataset enables us to develop accurate and efficient models, facilitating early detection and intervention in heart failure cases, thus improving patient outcomes, and reducing healthcare costs. In figure 1 is the output data distribution which shows that the classes are almost balanced.

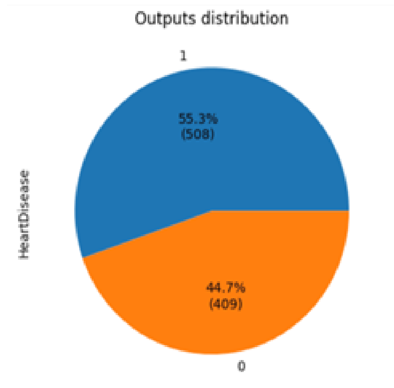

Figure 1: Output data distribution

B. Models:

We employed the Extra Trees Classifier algorithm for our classification task. The Extra Trees Classifier is an ensemble learning method that constructs multiple decision trees and combines their results to improve the overall accuracy and stability of the model. This method is particularly useful for handling high-dimensional data with complex feature interactions. To evaluate the performance of the classifier, we calculated various performance metrics, such as precision, recall, and F1-score.

The results from our classifier showed an overall accuracy of 0.83, with precision, recall, and F1-score values of 0.83 for both classes (0 and 1). The macro average and weighted average were also consistent across all performance metrics, indicating a balanced performance across both classes. These results suggest that the Extra Trees Classifier was able to effectively classify the given data with a relatively high degree of accuracy. However, feature selection using ant colony optimization achieved better performance.

C. Feature Selection:

The Genetic Algorithm (GA) is an evolutionary optimization technique inspired by the natural process of evolution. In our implementation, we used the

GeneticSelectionCV function with a population size of 50, crossover probability of 0.5, mutation probability of 0.2, and 100 generations. The fitness function was designed to maximize the weighted F1-score. GA selected the following features: ['Age', 'Sex', 'ChestPainType', 'RestingBP', 'FastingBS', 'RestingECG', 'MaxHR', 'ExerciseAngina', 'Oldpeak', 'ST_Slope']. The resulting classification report showed an accuracy of 0.83 and a weighted F1-score of 0.83.

Particle Swarm Optimization (PSO) is a population-based optimization method inspired by the social behavior of bird flocks or fish schools. In our implementation, the fitness function aimed to minimize the negative weighted F1score. We used a swarm size of 50 and a maximum of 100 iterations. PSO selected the following features: ['Sex', 'ChestPainType', 'RestingBP', 'Cholesterol', 'FastingBS', 'ExerciseAngina', 'Oldpeak', 'ST_Slope']. The classification report showed an improved accuracy of 0.90 and a weighted F1-score of 0.90.

Ant Colony Optimization (ACO) is an optimization technique inspired by the foraging behavior of ants. In our implementation, we used a fitness function like the one in PSO, with the goal of minimizing the negative weighted F1-score. We defined a search space based on the number of features and used a pheromone matrix initialized to ones. The algorithm was run for 100 iterations with 50 ants. ACO selected the following features: ['Age', 'Sex', 'ChestPainType', 'RestingBP', 'Cholesterol', 'FastingBS', 'RestingECG', 'MaxHR', 'ExerciseAngina', 'Oldpeak', 'ST_Slope']. The resulting classification report showed an accuracy of 0.84 and a weighted F1-score of 0.84.

D. Results:

The combination of Extra Trees Classifier and PSO for feature selection provided the best results in terms of accuracy and F1-score. This combination achieved an accuracy of 0.90 and a weighted F1-score of 0.90, outperforming the other methods. The selected features in this case were: ['Sex', 'ChestPainType', 'RestingBP', 'Cholesterol', 'FastingBS', 'ExerciseAngina', 'Oldpeak', 'ST_Slope']. This indicates that the Particle Swarm Optimization method successfully identified a subset of features that improved the classification performance in this specific problem.

Table 1 shows the weighted precision, weighted recall, weighted F1-score, and accuracy achieved by the classification model using the features selected by each method. Particle Swarm Optimization (PSO) provided the best performance across all metrics, with a weighted precision of 0.91, weighted recall of 0.90, weighted F1score of 0.90, and accuracy of 0.90.

Table 1: Results

Method	Precision	Recall	F1-score	Accuracy
Extra Trees (baseline)	0.83	0.83	0.83	0.83
Genetic Algorithm (GA)	0.83	0.83	0.83	0.83
Particle Swarm Optimization (PSO)	0.91	0.90	0.90	0.90
Ant Colony Optimization (ACO)	0.84	0.84	0.84	0.84

IV. DISCUSSION AND CONCLUSION

In conclusion, this research project demonstrated the effectiveness of using different feature selection techniques combined with the Extra Trees Classifier for classification tasks. Among the tested methods, Particle Swarm Optimization (PSO) outperformed the other techniques, achieving the highest weighted F1-score of 0.90 and accuracy of 0.90. The selected features by PSO were: ['Sex', 'ChestPainType', 'RestingBP', 'Cholesterol', 'FastingBS', 'ExerciseAngina', 'Oldpeak', 'ST_Slope']. This indicates that PSO was successful in identifying a subset of features that improved the classification performance. The results also show the importance of feature selection in enhancing model performance and reducing computational complexity. Future work for this project may involve testing additional feature selection methods and machine learning algorithms to further improve classification performance. Additionally, the use of advanced optimization algorithms, such as hybrid methods that combine the strengths of different techniques, could lead to more accurate and robust models. Moreover, incorporating domain knowledge to guide the feature selection process could help in identifying more informative features and potentially improve the model's interpretability. Lastly, exploring other evaluation metrics, such as AUC-ROC or Matthews Correlation Coefficient, could provide more insights into the performance of the classification model and help identify areas for improvement.

REFERENCES

[1] B. Soudan, F. F. Dandachi, and A. B. Nassif, "Attempting cardiac arrest prediction using artificial intelligence on vital signs from Electronic Health Records," *Smart Heal.*, vol. 25, p. 100294, Sep. 2022, doi: 10.1016/J.SMHL.2022.100294.

[1] C. W. Tsao et al., "Heart Disease and Stroke Statistics—2023 Update: A Report From the American Heart Association," Circulation, vol. 147, no. 8, pp. E93–E621, Feb. 2023, doi: 10.1161/CIR.0000000000001123.

[2] J. Xu, X. Luo, G. Wang, H. Gilmore, and A. Madabhushi, "A Deep Convolutional Neural Network for segmenting and classifying epithelial and stromal regions in histopathological images," Neurocomputing, vol. 191, pp. 214–223, May 2016, doi: 10.1016/j.neucom.2016.01.034.

[3] G. Motalleb, "Artificial neural network analysis in preclinical breast cancer," Cell J., vol. 15, no. 4, pp. 324–331, Dec. 2014, Accessed: Jan. 14, 2021. [Online]. Available: /pmc/articles/PMC3866536/?report=abstract

[4] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," Science (80-.)., vol. 349, no. 6245, pp. 255–260, Jul. 2015, doi: 10.1126/SCIENCE.AAA8415.

[5] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement Learning: A Survey," J. Artif. Intell. Res., vol. 4, pp. 237–285, May 1996, doi: 10.1613/JAIR.301.

[6] S. N. Sivanandam and S. N. Deepa, "Genetic Algorithm Optimization Problems," in Introduction to Genetic Algorithms, Springer Berlin Heidelberg, 2007, pp. 165–209. doi: 10.1007/978-3-540-73190-0_7.

[7] J. Kennedy and R. Eberhart, "Particle swarm optimization," Proc. ICNN'95 - Int. Conf. Neural Networks, vol. 4, pp. 1942–1948, doi: 10.1109/ICNN.1995.488968.

[8] N. A. Abujabal and A. B. Nassif, “Meta-heuristic algorithms-based feature selection for breast cancer diagnosis: A systematic review,” Int. Conf.
Electr. Comput. Commun. Mechatronics Eng.
ICECCME 2022, 2022, doi:
10.1109/ICECCME55909.2022.9988285.
[9] M. Dorigo, M. Birattari, and T. Stutzle, “Ant colony optimization,” IEEE Comput. Intell. Mag., vol. 1, no. 4, pp. 28–39, Nov. 2006, doi:
10.1109/MCI.2006.329691.
[10] M. Dorigo and G. Di Caro, “Ant colony optimization: A new meta-heuristic,” in Proceedings of the 1999 Congress on Evolutionary Computation, CEC 1999, IEEE Computer Society, 1999, pp. 1470–1477. doi:
10.1109/CEC.1999.782657.
[11] D. Karaboga and B. Basturk, “Artificial Bee Colony (ABC) optimization algorithm for solving constrained optimization problems,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif.
Intell. Lect. Notes Bioinformatics), vol. 4529 LNAI, pp. 789–798, 2007, doi: 10.1007/978-3540-72950-1_77/COVER.
[12] D. Karaboga, “Artificial bee colony algorithm,” Scholarpedia, vol. 5, no. 3, p. 6915, 2010, doi:
10.4249/scholarpedia.6915.
[13] T. M. Le, T. N. Pham, and S. V. T. Dao, “A Novel Wrapper-Based Feature Selection for Heart
Failure Prediction Using an Adaptive Particle
Swarm Grey Wolf Optimization,” Stud. Fuzziness Soft Comput., vol. 410, pp. 315–336, 2021, doi: 10.1007/978-3-030-70111-6_15/COVER.
[14] B. Soudan, F. F. Dandachi, and A. Bou Nassif, "Attempting cardiac arrest prediction using artificial intelligence on vital signs from electronic health records," Smart Health, vol. 25, pp. 100294, 2022.
[15] S. Irin Sherly and G. Mathivanan, “An efficient honey badger based Faster region CNN for chronc heart Failure prediction,” Biomed. Signal Process.
Control, vol. 79, p. 104165, Jan. 2023, doi:
10.1016/J.BSPC.2022.104165.
[16] D. Chicco and L. Oneto, “An Enhanced Random Forests Approach to Predict Heart Failure from Small Imbalanced Gene Expression Data,” IEEE/ACM Trans. Comput. Biol. Bioinforma., vol.
18, no. 6, pp. 2759–2765, 2021, doi:
10.1109/TCBB.2020.3041527.
[17] S. J. Khiabani, A. Batani, and E. Khanmohammadi, “A hybrid decision support system for heart failure diagnosis using neural networks and statistical process control,” Healthc.
Anal., vol. 2, no. September, p. 100110, 2022, doi:
10.1016/j.health.2022.100110.
[18] M. M. Yaqoob, M. Nazir, M. A. Khan, S. Qureshi, and A. Al-Rasheed, “Hybrid Classifier-Based Federated Learning in Health Service Providers for Cardiovascular Disease Prediction,” Appl. Sci., vol. 13, no. 3, 2023, doi: 10.3390/app13031911.
[19] P. K. Sahoo and P. Jeripothula, “Heart Failure Prediction Using Machine Learning Techniques,” SSRN Electron. J., Dec. 2020, doi:
10.2139/SSRN.3759562.
autoencoders,” Measurement, vol. 205, p. 112166,
Dec. 2022, doi:
10.1016/J.MEASUREMENT.2022.112166
[20] A. Noor, L. Ali, H. T. Rauf, U. Tariq, and S. Aslam, “An integrated decision support system for heart failure prediction based on feature
transformation using grid of stacked