2024 Asian Conference on Intelligent Technologies (ACOIT)
Karnataka, India. Sep 6-7, 2024

Smart Heart: Leveraging SVM for Early and Accurate Heart Failure Detection

Goldy Verma[1]
Chitkara University Institute of Engineering and Technology, Chitkara University,
Punjab, India
goldy.verma@chitkara.edu.in

Kanwarpartap Singh Gill[2]
Chitkara University Institute of Engineering and Technology, Chitkara University,
Punjab, India
kanwarpartap.gill@chitkara.edu.in

Mukesh Kumar[3]
Computer Science & Engineering, Graphic Era Hill University,
Dehradun, Uttarakhand, India, 248002
kumarsidcse@gmail.com

Ruchira Rawat[4]
Computer Science & Engineering, Graphic Era Deemed to be University,
Dehradun, Uttarakhand, India, 248002
ruchira.rawat.cse@geu.ac.in

2024 Asian Conference on Intelligent Technologies (ACOIT) | 979-8-3503-7495-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/ACOIT62457.2024.10941680

***Abstract*— In this study, the predictive modeling employed is the support vector machine (SVM) method, which is adept at handling extensive and complex health data. Despite the high-stress lifestyles many of us lead today, it seems unlikely that heart issues were as prevalent in the past. Heart failure occurs when the heart no longer functions properly. SVM helps identify individuals at risk of serious vascular problems by predicting potential health issues. The SVM demonstrates an impressive 79% accuracy in classifying various clinical indicators and risk factors, which is notable. This suggests that, if utilized effectively, such predictive tools could significantly transform healthcare by detecting heart conditions early and potentially reducing costs. The study emphasizes the importance of advanced, informed computational methods in improving early detection and treatment of heart failure, ultimately leading to better health monitoring strategies.**

Keywords— Machine Learning, Heart Failure, Support Vector machine, Efficacy, Health, Mental Health, Developing countries, Classification Analysis

I. INTRODUCTION

In today's fast-paced world, people often find themselves so engrossed in daily responsibilities that they have little time to focus on their health and self-care. This relentless drive to meet obligations frequently results in both mental and physical stress, making it difficult to maintain good health. As a result, cardiovascular diseases like heart failure have become significant challenges to both public health and personal well-being. Heart failure, a serious condition where the heart struggles to pump sufficient blood to meet the body's needs, has become a major global health concern. It can be caused by weakened heart muscles or poor pumping capacity, leading to various symptoms that significantly reduce patients' lifespan and quality of life. Despite advances in medical research that have led to treatments improving symptoms and extending life expectancy, heart failure continues to have a considerable impact. Recognizing the need for early detection and intervention, this study aims to predict and reduce the risk of heart failure using modern computational techniques. Specifically, it employs support vector machines (SVM), a powerful machine learning method known for its effectiveness in classification tasks. SVMs offer robust analytical capabilities with their supervised learning framework and the concept of maximizing margins to detect patterns in large, complex datasets. In this study, SVM is used to analyze a dataset with various clinical indicators and risk factors. The advantage of SVM lies in its memory efficiency and effectiveness in high-dimensional scenarios. The study demonstrates that SVM can achieve an impressive accuracy rate of 79% in identifying heart failure risk factors. This paper highlights the potential of SVM in cardiovascular risk assessment, contributing to the expanding field of predictive analytics in healthcare. By improving the accuracy of risk identification, this work provides valuable insights for developing precise prediction models for clinical use.

II. LITERATURE

There has been a growing interest in predicting heart failure (HF) to revolutionize healthcare delivery and improve patient outcomes. Extensive research has focused on the capabilities of machine learning algorithms, particularly support vector machines (SVMs), in identifying individuals at risk of heart failure. Advanced machine learning (ML) techniques have been employed to enhance patient outcomes and diagnostic accuracy, making the prediction and management of cardiac diseases a critical research area. Hoque et al. [1] illustrate the robustness of SVMs in handling large datasets, demonstrating their effectiveness in heart disease prediction. Similarly, Ahmed et al. [2] explore the use of various classification algorithms for predicting heart failure patient survival, emphasizing their value in clinical decision-making. Kumar et al. [3] highlight the increasing importance of integrating digital health resources into predictive modeling, presenting an advanced ML-based framework for early detection of heart failure using mobile health (mHealth) technologies. Victor and colleagues [4] show that non-invasive techniques for ML algorithm-based heart failure assessment can reduce patient pain and increase compliance. Elsedimy et al. [5] provide a comprehensive examination of multiple machine learning techniques for heart disease prediction, offering a valuable resource for researchers in this field. Harish and Sabitha [6] demonstrate significant improvements in prediction accuracy with SVM and a novel penalty-based logistic regression classifier for cardiac disease prediction. Mansur Huang et al. [7] underscore the importance of early intervention by applying ML techniques for early heart failure prediction to improve patient outcomes. Prasetyo et al. [9] compare two heart failure prediction algorithms, logistic regression and SVM, to determine the optimal method for accurate predictions. Behera [12] discusses the reliability and accuracy of SVM in various research contexts for heart disease prediction. These studies collectively highlight the essential role of machine learning in enhancing the detection and treatment of cardiac diseases. The use of SVM, logistic regression, and other ensemble methods aims to improve diagnostic accuracy, facilitate early detection, and ultimately enhance patient care. As the healthcare landscape continues to evolve, integrating these advanced predictive models with existing clinical practices will be crucial in addressing the increasing burden of heart disease worldwide. With ongoing research and development, machine learning holds great promise for transforming heart disease treatment and improving patient outcomes.

III. INPUT DATASET

Heart failure frequently results from cardiovascular diseases (CVD). This dataset includes 12 features that can help predict mortality rates associated with heart failure. Preventative treatments targeting behavioral risk factors such as smoking, poor diet, obesity, lack of physical activity, and excessive alcohol consumption can avert most cardiovascular diseases. Early diagnosis and management of cardiovascular disease or individuals at high cardiovascular risk (due to factors like hypertension, diabetes, high cholesterol, or existing conditions) are crucial. A machine learning model can be particularly useful in this area. The dataset, sourced from the open-access Kaggle platform, includes features such as anemia, diabetes, ejection fraction, high blood pressure, platelets, serum sodium and creatinine levels, sex, smoking status, time, and death events. The final feature, death event, is coded as 0 for no heart failure and 1 for heart failure. Table 1 below displays the dataset with various parameters.

TABLE I. DIFFERENT PARAMETERS DESCRIBING THE HEART FAILURE.

age	anaemia	creatinine	diabetes	ejection	High blood pressure	sex
75	0	582	0	20	1	1
55	0	7861	0	38	0	1
65	0	146	0	20	0	1
50	1	111	0	20	0	1
65	1	160	1	40	0	0
90	1	47	0	60	1	1
70	0	92	0	40	1	0
42	0	102	1	38	0	1

IV. PROPOSED METHODOLOGY

This paper introduces a supervised max margin model utilizing the support vector machine (SVM) for regression and classification of data. Applications include image categorization, spam detection, text and hypertext processing, and managing high-dimensional data. The primary objective of this approach is to identify the optimal hyperplane in an N-dimensional space that can segregate data points into different classes within the feature space. There are two types of SVMs: linear (or simple) vector machines and non-linear (or kernel) vector machines. The linear SVM uses a linear decision boundary to separate data points into various classes, whereas the non-linear SVM is employed when data cannot be separated by a straight line, utilizing a non-linear boundary instead. SVMs are particularly advantageous due to their effectiveness in high-dimensional environments and their efficient memory usage. They are widely used in healthcare for tasks such as disease prediction, signal processing, natural language processing, and image and audio recognition. Figure 1 illustrates the proposed methodology for the SVM.

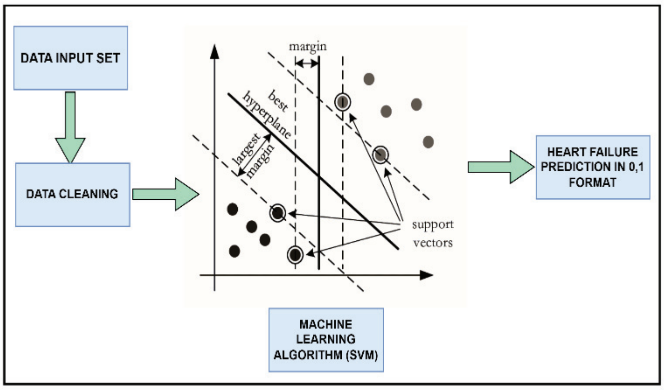

Fig. 1. Proposed methodology for the heart failure prediction using SVM

V. RESULTS

A. Confusion Matrix Analysis

This theoretical framework evaluates the prediction performance of a binary classification model using a confusion matrix with values [(0.57, 0.1)] and [(0.11, 0.22)]. The diagonal elements represent the correctly identified cases: true positives (TP) at 0.22 and true negatives (TN) at 0.57, shown in the upper-left cell. Conversely, the off-diagonal elements indicate misclassifications: false positives (FP) at 0.1 in the upper right cell and false negatives (FN) at 0.11 in the lower-left cell. These values demonstrate the model's ability to differentiate between classes, with higher true positive rates indicating better predictive performance, while higher false positive and false negative rates suggest areas for improving generalization and classification accuracy.

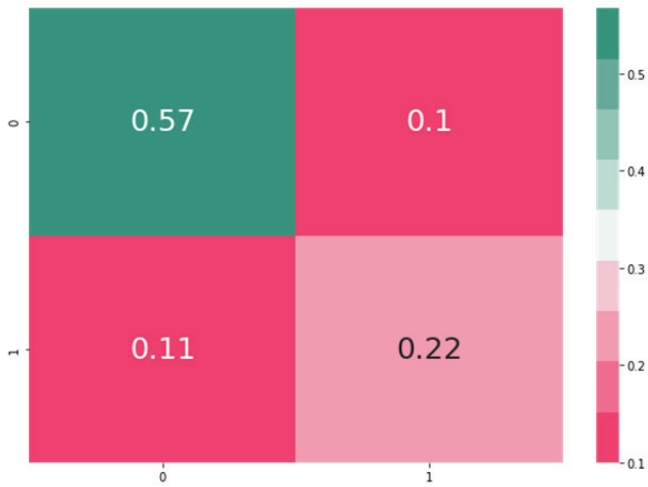

Fig. 2. Confusion Matrix Analysis

B. Classification Report Analysis

The table summarizes the performance of a machine learning model on a categorization task, presenting accuracy, macro average, precision, recall, F1-score, and support for two classes (0 and 1). Precision indicates the proportion of positive predictions that were correct. High accuracy means the model correctly predicted most positive cases. This model showed strong performance in recognizing positive instances, with precision scores of 0.84 and 0.69 for the two classes. Recall shows how well the model identified actual positive cases. A high recall means the model captured most relevant positive instances, which it did with scores of 0.85 and 0.67 for both classes. The F1 Score, a harmonic mean of recall and precision, provides a balanced measure. The high F1 scores of 0.84 and 0.68 for both classes align with the precision and recall results. Support denotes the number of data points in each class, with 60 data points in class 0 and 30 in class 1. Accuracy represents the proportion of correct predictions out of the total instances, with an overall accuracy of 0.76. The macro average, the unweighted mean of recall and precision for each class, is also 0.76, offering a standardized performance measure. Weighted averages, considering class distribution, indicate a slightly better performance for the class with more examples (class 0), with a weighted precision, recall, and F1-score of 0.79. Overall, the table shows that the model performed well in this categorization task, achieving respectable accuracy, recall, and F1 scores for both classes, and demonstrating fairly high precision.

TABLE II. CLASSIFICATION REPORT ANALYSIS

	precision	recall	F1-score	Support
0	0.84	0.85	0.84	60
1	0.69	0.67	0.68	30
Accuracy			0.79	90
Macro avg	0.76	0.76	0.76	90
Weighted avg	0.79	0.79	0.79	90

VI. CONCLUSION

This research explores the application of support vector machines (SVM) on high-dimensional clinical datasets to predict heart failure, leveraging its effective classification capabilities. The incidence of cardiovascular diseases, including heart failure, has risen due to the modern lifestyle's focus on work and neglect of personal health. Therefore, accurately predicting and mitigating the risk of heart failure is crucial for improving patient outcomes and reducing healthcare costs. Our study examined a comprehensive dataset encompassing various clinical characteristics and risk factors associated with heart disease. Through meticulous feature selection, model optimization, and evaluation, we achieved a notable prediction accuracy rate of 79%, underscoring SVM's potential as a reliable tool for risk assessment in clinical settings. The literature review illuminated the current state of heart failure prediction, demonstrating the increasing popularity of machine learning methods, particularly SVM, for accurate risk classification. Our findings emphasize the significance of advanced computational techniques like SVM in the early diagnosis and management of heart failure. Future advancements in machine learning could enhance prediction accuracy and enable more personalized treatments, ultimately improving patient outcomes and advancing precision medicine in cardiovascular care.

REFERENCES

[1] Hoque, R., Billah, M., Debnath, A., Hossain, S.S. and Sharif, N.B., 2024. Heart Disease Prediction using SVM. *International Journal of Science and Research Archive*, *11*(2), pp.412-420..

[2] Ahmed, R., Bibi, M. and Syed, S., 2023. Improving heart disease prediction accuracy using a hybrid machine learning approach: A comparative study of svm and knn algorithms. *International Journal of Computations, Information and Manufacturing (IJCIM)*, *3*(1), pp.49-54.

[3] Kumar, D., Balraj, K., Seth, S., Vashista, S., Ramteke, M. and Rathore, A.S., 2024. An improved machine learning-based prediction framework for early detection of events in heart failure patients using mHealth. *Health and Technology*, pp.1-18.

[4] Victor, O.A., Chen, Y. and Ding, X., 2024. Non-Invasive Heart Failure Evaluation Using Machine Learning Algorithms. *Sensors*, *24*(7), p.2248.

[5] Elsedimy, E.I., AboHashish, S.M. and Algarni, F., 2024. New cardiovascular disease prediction approach using support vector machine and quantum-behaved particle swarm optimization. *Multimedia Tools and Applications*, *83*(8), pp.23901-23928.

[6] Harish, P. and Sabitha, R., 2024, February. Analyzing the efficiency of heart disease prediction using SVM and an innovative penalty based logistic regression classifier (IPLR). In *AIP Conference Proceedings* (Vol. 2729, No. 1). AIP Publishing.

[7] Mansur Huang, N.S., Ibrahim, Z. and Mat Diah, N., 2021. Machine learning techniques for early heart failure prediction. *Malaysian Journal of Computing (MJoC)*, *6*(2), pp.872-884..

[8] Gill, K.S., Sharma, A., Anand, V. and Gupta, S., 2022, December. An Automated Coronary Artery Disease Diagnosis System using Machine Learning. In 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 1173-1178). IEEE.

[9] Prasetyo, A., Mulia, D.A. and Octavina, K.K., 2023. Comparative Study of Heart Failure Prediction Algorithm: Logistic Regression and SVM. *International Journal of Intelligent Systems and Applications in Engineering*, *11*(2), pp.518-522.

[10] Gill, K.S., Anand, V. and Gupta, R., 2023, July. Arrhythmia Classification Using ECG Image Dataset Using Machine Learning Approach on DenseNet121 Model. In 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-4). IEEE.

[11] Behera, M.P., Sarangi, A., Mishra, D. and Sarangi, S.K., 2023. A hybrid machine learning algorithm for heart and liver disease prediction using modified particle swarm optimization with support vector machine. *Procedia Computer Science*, *218*, pp.818-827.