Heart Failure Prediction using Support Vector Machine

Christian Joshua Alberto
College of Computing and Information Technologies
National University - Manila
Manila, Philippines
albertocq@students.national-u.edu.ph.

Mark Rhey Anthony De luna
College of Computing and Information Technologies
National University - Manila
Manila, Philippines
delunamd@students.national-u.edu.ph

Rodney Lei Estrada
College of Computing and Information Technologies
National University - Manila
Manila, Philippines
estradard@students.national-u.edu.ph

***Abstract*—The research investigates the use of Support Vector Machines (SVM) to predict heart failure outcomes based on 12 clinical attributes from 299 patient records. Employing the Radial Basis Function kernel, the SVM model classifies patients into risk categories for heart failure. The study examines model performance and the effect of hyperparameter tuning techniques, including GridSearchCV, Optuna, and Hyperopt. Findings reveal that the default SVM configuration achieves satisfactory results on structured data, while tuning enhances performance for more complex datasets. The model shows potential for early diagnosis, aiding healthcare providers in predictive decision-making for heart failure management.**

I. Introduction

Heart failure poses a large health challenge on a global scale, leading to considerable illness and death across the world. Heart failure is acknowledged as a leading reason for hospitalization, especially among elderly individuals, affecting over 64 million people [1]. The worldwide impact of Heart Failure is on the rise due to growing older populations and behaviors like lack of physical activity and poor diets, which raise the chances of heart disease. By 2030, it is estimated that the expenses linked to Heart Failure in the United States will increase by almost 127%, causing significant pressure on healthcare systems [2]. These patterns highlight the critical requirement for creative and dependable techniques for detecting illnesses early, allowing for prompt interventions to lower hospital re-admissions and enhance patient results.

Traditional diagnostic methods often lack the precision required for early detection of Heart Failure in patients at high risk, creating a gap in preventive care. To address this challenge, researchers are increasingly turning to machine learning (ML) techniques, which hold promise for enhancing diagnostic accuracy. Among various ML approaches, Support Vector Machines (SVM) have shown particular effectiveness in high-dimensional classification tasks, making them suitable for medical data where multiple clinical attributes can influence the prediction. SVM models are especially advantageous due to their ability to handle complex datasets, manage noise, and provide robust classification, which is essential in the nuanced environment of medical diagnosis [3].

This study focuses on applying SVM to predict Heart Failure outcomes based on 12 key clinical attributes from a dataset of 299 patient records, including variables such as ejection fraction, serum creatinine, and age. By employing the Radial Basis Function (RBF) kernel, the model categorizes patients into high- and low-risk groups, enabling targeted intervention. Furthermore, the study explores advanced hyperparameter tuning methods—GridSearchCV, Optuna, and Hyperopt—to enhance the model's predictive performance, particularly on more complex datasets. These tuning techniques allow for fine-tuning the model's parameters, thus optimizing accuracy and robustness.

Initial findings demonstrate that while the default SVM configuration yields satisfactory results with structured clinical data, tuning enhances its predictive performance, making the model more adaptable to diverse patient profiles. With its improved classification capabilities, the SVM model offers substantial potential for aiding healthcare providers in early Heart Failure diagnosis, which could ultimately lead to better patient outcomes, fewer hospital readmissions, and reduced healthcare costs. This approach exemplifies how machine learning can bridge existing diagnostic gaps, transforming Heart Failure care from a reactive to a proactive approach.

II. Review of Related Literature

1) **Overview of Key Concepts and Background Information**
Relevant Concepts, Models, and Methods:
Heart failure(HF) prediction has traditionally relied on statistical models such as logistic regression and decision trees, which classify patients based on clinical features like blood pressure, age, and serum creatinine. These models, however, often lack the flexibility to

capture high-dimensional and nonlinear relationships typical in complex medical data [4]. Support Vector Machines (SVM) have emerged as effective ML models for handling binary classification problems, particularly with non-linear kernels like the Radial Basis Function (RBF) [5]. Key methods in SVM-based classification include hyperparameter tuning techniques like GridSearchCV, Optuna, and Hyperopt, which enhance accuracy by optimizing parameters [6].

Historical Development:
The prediction of heart disease risk has evolved from simple scoring systems, such as the Framingham Risk Score, to complex ML algorithms that leverage non-linear models and big data. Earlier approaches applied manually curated clinical features in linear models, while recent methods detect subtle patterns among numerous variables [7]. The shift from linear models to kernel-based SVMs marked a significant improvement in capturing the nuanced relationships in cardiovascular data. This progression directly informs the current study, as SVM offers a robust alternative to traditional methods by efficiently handling complex, nonlinear relationships in clinical data [8].

2) **Review of Relevant Research Papers**
Relevant Studies:

- *Improving Risk Prediction in Heart Failure (MARKER-HF Model)*: This study applied a Boosted Decision Tree (BDT) algorithm to predict HF mortality, achieving an AUC of 0.88 by focusing on significant clinical variables like diastolic blood pressure and serum creatinine. The model's emphasis on biomarkers is aligned with the current study's goal of refining HF risk prediction [9].
- *Machine Learning-Based Prediction of Survival in HF Patients*: Using WHO classifications for HF, this study explored various models, including Decision Trees, Random Forest (RF), and XGBoost, for HF survival prediction. Highlighting challenges in capturing complex interactions, this research supports the advantage of advanced ML techniques for HF outcomes [10].
- *Comparative Study on HF Prediction Models*: This study evaluated 18 ML models, including logistic regression, SVM, and ensemble methods, demonstrating that z-score normalization combined with SMOTE improved model robustness and accuracy. These findings reinforce the current research's focus on SVM with tuning techniques to enhance predictive accuracy [11].

Relation to Current Study:
These studies both inform and contrast with the current research. Unlike prior studies that focus on ensemble models, this research investigates the potential of a tuned SVM model to achieve high accuracy on structured clinical datasets without needing ensemble methods [11].

3) **Current State of the Art**
Best-Performing Non-ML and ML Techniques:
Currently, logistic regression and decision trees remain prevalent in clinical settings for their interpretability. In contrast, ML techniques like SVM, RF, and XGBoost are increasingly preferred for their flexibility and ability to handle complex datasets [12]. State-of-the-art methods in the field typically integrate ensemble learning and parameter tuning to achieve optimal performance.

Advantages and Limitations:
Traditional models are limited in their ability to manage high-dimensional data, whereas SVM, especially with RBF kernels, captures nonlinear relationships and improves prediction accuracy. Nonetheless, SVMs are computationally intensive, and careful parameter tuning is often essential to achieve peak performance, a limitation this study addresses through methods such as GridSearchCV, Optuna, and Hyperopt [6].

4) **Prior Attempts to Solve the Same Problem**
Prior Attempts to Solve the Same Problem:
Notable contributions to HF prediction include the MAGGIC risk score, the MARKER-HF model, and recent comparative studies on ML-based HF models. These studies lay the groundwork for advanced HF prediction models, though proprietary models from companies like IBM and Philips often lack transparency and generalizability [4], [9]

Successes and Shortcomings:
Previous models have successfully integrated clinical data to predict HF, yet many still struggle with patient variability and intricate variable relationships. Models like the MAGGIC risk score and proprietary AI models either require extensive testing or lack generalizability, motivating the need for adaptable, high-performance ML solutions like the tuned SVM model explored in this study [8], [12].

Summary of Key Questions

1) **What research has already been done on this topic?**
HF prediction research includes a progression from simple risk scores to complex ML models, with recent emphasis on ML methods such as SVM and ensemble learning for improved prediction accuracy [4], [7], [9].
2) **How do existing studies relate to or differ from your research?**
Previous studies provide a foundational framework yet often rely on ensemble techniques or specific biomark-

ers. This research aims to optimize a single SVM model using hyperparameter tuning, demonstrating that a well-tuned SVM can provide competitive accuracy without ensemble complexity [10], [11].

3) **What are the key methods, theories, and models that inform your work?**
Central to this work are SVM with RBF kernel, hyperparameter tuning methods like GridSearchCV, Optuna, and Hyperopt, and the focus on clinical data for feature selection [6].
4) **What gaps exist in the current literature that your research fills?**
This research fills the gap in optimized, high-performance SVM models for HF risk prediction that don't depend on extensive clinical testing or complex ensemble structures [12].
5) **How does your proposed approach contribute to advancing the field?**
By demonstrating the potential of a fine-tuned SVM model, this study provides a streamlined, accurate approach for early HF diagnosis and clinical decision-making support [7], [8].

III. METHODOLOGY

A. Data Collection

The researchers used a publicly available dataset from Kaggle. The dataset includes a collection of medical records for 299 heart failure patients which consisted of 105 women and 194 men with relevant features like age, blood pressure, heart rate, ejection fraction, and other clinical parameters.

The data was originally collected by Davide Chicco and Giuseppe Jurman at the Faisalabad Institute of Cardiology and at the Allied Hospital in Faisalabad (Punjab, Pakistan), during April - December 2015. The data was gathered through standardized clinical assessments, ensuring the consistency in feature measurements. [13]

B. Data Pre-Processing

In data cleaning, missing values were identified and handled based on clinical relevance. The researchers imputed missing values for features with a few missing entries using the mean. Outliers in clinical features were identified using statistical threshold, and these are removed to maintain data integrity.

To ensure an even distribution of the target variable, the Synthetic Minority Over-sampling Technique (SMOTE) was used to balance out the value count of the death event.

Clinical features such as blood pressure, ejection fraction, serum creatinine, and age varied greatly in scale, which could affect the SVM's performance. To address this, the researchers used Standard Scalling to normalize features, transforming each value to a range of 0 to 1. Categorical features, such as gender, were encoded using one-hot encoding to convert them into a suitable format for SVM.

C. Experimental Setup

To ensure efficient dependency management and reproducibility, the researchers used Python Poetry to manage and maintain the project environment. Poetry streamlines the installation of libraries and handles project dependencies, version control, and virtual environments effectively

The following tools and libraries were utilized to implement, train, and evaluate the Support Vector Machine (SVM) model for heart failure prediction:

- **Python**: The primary programming language used for the entirety of the analysis and model development due to its versatility and extensive library support.
- **NumPy**: essential for numerical computations, such as matrix operations, required in model training and data transformation.
- **Pandas**: Employed for data manipulation and analysis, including
- **Scikit-learn**: used for implementing the SVM model and providing tooles for data scaling, hyperparameter tuning (via GridSearchCV), and model evaluation.
- **Imbalance-learn** (imblearn): specifically used Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance by generating synthetic examples, thereby enhancing the model's ability to predict under-represented cases.
- **Matplotlib** and **Seaborn**: Visualization libraries used for data evaluation, including generating feature distribution plots, correlation heatmaps, and confusion matrices for performance evaluation.

To build a reliable predictive model, the data was divided into distinct training, validation and test sets to maintain independence in evaluation and ensure reliability. The training set comprising 80% of the data was used to fit the model, allowing it to learn patterns associated with heart failure. A validation set is also created consisting 20% of the data, used for hyperparameter tuning through cross-validation. Lastly, the test set representing 20% of the data, was held out for the final evaluation to simulate real-world model performance of the data.

This data split was performed using train_test_split from scikit-learn, ensuring a balanced distribution of the target variable in each subset. Additionally, class balancing was implemented using SMOTE from the imbalance-learn library. SMOTE addressed the class imbalance by synthetic example generation.

Several hyperparameters were tuned to optimize the SVM model's performance, with GridSearchCV used for an exhaustive search over specified parameter values:

- Kernel Type(kernel): Three types of kernels—linear, radial basis function (RBF), and polynomial—were evaluated to determine the most appropriate kernel for the heart failure dataset. The RBF kernel is generally effective for non-linear relationships, which are common in clinical data, while linear and polynomial kernels were also considered to capture simpler or polynomial relationships.

- Regularization Parameter(C): T: This hyperparameter controls the trade-off between achieving a low training error and maximizing the decision boundary margin. Higher values of C led to narrower margins with fewer misclassifications, while lower values allowed wider margins but tolerated more misclassifications. The range for C was systematically tested to identify the optimal balance.
- Gamma: Relevant specifically for the RBF kernel, gamma determines the influence range of each data point. Lower values of gamma imply a broader influence for each data point, resulting in smoother decision boundaries, while higher values create more complex boundaries with localized influence. Tuning gamma was essential to capture the most meaningful decision boundary for heart failure prediction.
- Cross-Validation Folds (CV): A 5-fold cross-validation setting was applied to evaluate the hyperparameter configurations. Each fold splits the data into five equal parts, with each part used once as the validation set and the remaining four as the training set. This ensured that the model was validated across all data partitions and mitigated potential overfitting.

D. SVM Algorithm

Supervised learning method for classification known as the Support Vector Machine (SVM) is one of the most effective algorithms used for heart failure prediction. SVM utilizes a linear decision boundary to identify the normal vector in the feature space, aiming to classify data into two sets: for instance, patients with heart failure and those without. In two-dimensional scenarios, this separation is represented by a line, while in higher dimensions, it is represented by a hyperplane. The principle behind SVM is to maximize the margin between the closest data points (support vectors) from each class, ensuring that the classification is robust and not overly sensitive to boundary points [14]. The decision boundary is mathematically represented as:

$$\beta_1 + \beta_1 * x_1 + \beta_2 * x_2 + ... + \beta_n * x_n = 0 \tag{1}$$

Equation 1: Hyperplane Representation

- β_0: The **intercept** or bias term
- $\beta_1, \beta_2, ..., \beta_n$: The **coefficients** (weights) for each feature in the dataset.
- $x_1, x_2, ..., x_n$: The **feature values** (e.g., age, serum creatinine, ejection fraction) for each data point.

SVM algorithms incorporate the kernel trick to map new data that is irreversibly nonlinear to a higher-dimensional space to easily find a linear separation. For the purpose of this work, the Radial Basis Function (RBF) kernel is applied; this kernel is suitable for cases of a non-linear mapping of the clinical features such as age, ejection fraction and serum creatinine, to the corresponding outcomes.

SVM algorithm's advantages in this project is that the algorithm is effective in high-dimensional space, secondly, the feature of C overfitting, third it uses only the support vectors instead of the whole data space [15]. The second factor is that the choice of the regularization parameter balances between having a smooth decision boundary and the number of misclassified data in the training data set. This makes SVM particularly suitable for predicting heart failure, so that the model, avoiding overfitting, will generalize well to new, unseen data of patients in the field of medical applications for early detection and diagnosis of the disease.

IV. RESULTS AND DISCUSSIONS

The research evaluates Support Vector Machine (SVM) model performance in predicting heart failure outcomes under different configurations: hyperparameter tuning, feature scaling, and data balancing techniques, including SMOTE and undersampling. The following sections discuss how each configuration impacts model performance, using metrics such as accuracy, F1-score, precision, and recall as benchmarks.

1) Sampling Manipulation Techniques

Table I. Sampling Manipulation comparison using Synthetic Minority Oversampling Technique (SMOTE), and Undersampling trained on simple SVM.

SM	*Accuracy*	*Precision*	*Recall*	*F1-score*
No Manipulation	0.75	0.78	0.75	0.73
Undersampling	0.85	0.85	0.84	0.85
SMOTE	0.84	0.84	0.84	0.84

Table I shows the results from the experimentation conducted on the SVM model. The highest F1-score was achieved using the undersampling technique, which improved the balance between precision and recall. Undersampling not only enhanced accuracy but also produced a balanced classification report, indicating that this technique effectively improved the model's ability to predict both classes accurately. However, the researcher chose to utilize the Synthetic Minority Oversampling Technique (SMOTE) for its ability to create synthetic samples rather than removing data points, as the way data was gathered did not involve duplicate values. This approach helps maintain the integrity of the dataset while addressing class imbalance. While SMOTE also performed well, undersampling yielded the best overall metrics in this scenario, demonstrating a trade-off between accuracy and data preservation.

2) Feature Scaling

Table II. Feature Scaling comparison using without scaling, MinMax Scaler, and Standard Scaler trained on simple SVM.

Feature Scaling	*Accuracy*	*Precision*	*Recall*	*F1-score*
No Scaling	0.58	0.34	0.58	0.43
Standard Scaler	0.75	0.78	0.75	0.73
SMOTE	0.70	0.73	0.70	0.67

Table II illustrates the impact of feature scaling on model performance. The StandardScaler consistently produced the best results, significantly outperforming the no-scaling condition. The model's precision and recall indicate that without scaling, the model struggles to generalize, leading to lower overall effectiveness. This emphasizes the importance of feature scaling in improving model accuracy and robustness.

3) Hyperparameter Tuning

Table III. Hyperparameter Tuning Cross Validation comparison using GridSearchCV, Optuna, and Hyperopt. Trained under SMOTE manipulated data. .

HT	*Accuracy*	*Precision*	*Recall*	*F1-score*
GridSearchCV	0.84	0.84	0.84	0.84
Optuna	0.80	0.81	0.80	0.80
Hyperopt	0.83	0.83	0.83	0.83

Table III summarizes the results of hyperparameter tuning using various optimization methods, including GridSearchCV, Optuna, and Hyperopt, all trained on SMOTE-manipulated data. Both GridSearchCV and Hyperopt achieved the highest performance metrics, each with an accuracy, precision, recall, and F1-score of 0.84, indicating their effectiveness in optimizing the SVM model under the given conditions. This consistency across metrics suggests that these methods can effectively balance model performance and classification accuracy.

However, it's crucial to note that higher accuracy does not always imply better performance in all classification metrics. The SVM model's propensity for generating false positives highlights the importance of context in evaluation. In medical diagnostics, identifying false positives (misclassifying patients as having heart failure when they do not) may be more acceptable than failing to detect true positives (failing to identify actual heart failure cases). Thus, the ramifications of false negatives can be particularly severe in this field. But the researcher chose to use the GridSearchCV as we do metrics now for aiming accuracy.

V. Conclusions

The research investigates the early prediction of heart failure outcomes using Support Vector Machines (SVM) on clinical data collected from 299 heart failure patients. By analyzing 12 attributes, including age, serum creatinine, and ejection fraction, the study assesses the efficacy of SVM in classifying survival outcomes. Additionally, it explores the impact of hyperparameter tuning through methods such as GridSearchCV, Optuna, and Hyperopt. The results indicate that simple hyperparameter tuning using GridSearchCV yields slightly better performance than more complex tuning methods, suggesting that default settings may suffice for datasets with structured simplicity. Nevertheless, hyperparameter tuning has proven valuable for enhancing performance in more complex datasets.

Key findings highlight the advantages of employing the Synthetic Minority Over-sampling Technique (SMOTE) over undersampling. While undersampling yielded slightly better performance by a narrow margin, the researcher chose to use SMOTE due to its ability to generate additional synthetic data points through interpolation rather than duplicating existing samples. This approach effectively addresses class imbalance without discarding valuable data, which can diminish predictive accuracy. Regarding feature scaling, the use of the StandardScaler consistently delivered the best performance, closely followed by MinMax scaling, while models trained without any scaling performed poorly. This underscores the critical role of scaling in maximizing SVM accuracy.

These results illustrate SVM's potential in predicting heart failure risks, empowering healthcare providers to identify high-risk patients early. The model has promising applications in predictive healthcare systems, assisting clinical decision-making and facilitating early interventions.

However, the study's limitations, including the relatively small sample size and highly structured data, restrict the generalizability of the findings. Future research should aim to expand the sample size, incorporate additional clinical features, and compare SVM with other machine learning models across varied datasets to enhance accuracy. This study emphasizes the significance of understanding the metrics of classification reports, as not all models with the highest accuracy are necessarily the best. It serves as a foundation for further exploration into SVM's utility in healthcare predictive systems and contributes to the broader application of machine learning in healthcare settings.

References

[1] B. K. Lee, "Global burden of heart failure," European Heart Journal, vol. 44, Supp. 2, pp. 655-876, 2023. https://academic.oup.com/eurheartj/article/44/Supplement_2/ehad655.876/7391989.

[2] D. D. Ezekowitz et al., "Global public health burden of heart failure: an updated review," Cardiac Failure Review, 2023. https://www.cfrjournal.com/articles/global-public-health-burden-heart-failure-updated-review

[3] A. Z. Ali and M. R. Mahmud, "Support vector machine in medical diagnosis," IEEE Access, vol. 8, pp. 123–134, 2022.

[4] Bitton, A., & Gaziano, T. (2010). The Framingham Heart Study's impact on global risk assessment. Progress in Cardiovascular Diseases, 53(1), 68–78. https://doi.org/10.1016/j.pcad.2010.04.001

[5] Sunday Adeola Ajagbe. (2022). Support Vector Machines (SVMS) based advanced healthcare system using machine learning techniques. International Journal of Innovative Research in Computer and Communication Engineering, 3007–3008. https://doi.org/10.15680/IJIRCCE.2022.1005020

[6] Vertes, D. (2021, December 16). Beyond Grid Search: Hypercharge Hyperparameter Tuning for XGBoost. Medium. https://towardsdatascience.com/beyond-grid-search-hypercharge-hyperparameter-tuning-for-xgboost-7c78f7a2929d

[7] Nakai, M., Watanabe, M., Kokubo, Y., Nishimura, K., Higashiyama, A., Takegami, M., Nakao, Y. M., Okamura, T., & Miyamoto, Y. (2020). Development of a Cardiovascular Disease Risk Prediction Model Using the Suita Study, a Population-Based Prospective Cohort Study in Japan. Journal of Atherosclerosis and Thrombosis, 27(11), 1160–1175. https://doi.org/10.5551/jat.48843

[8] Javaid, A., Zghyer, F., Kim, C., Spaulding, E. M., Isakadze, N., Ding, J., Kargillis, D., Gao, Y., Rahman, F., Brown, D. E., Saria, S., Martin, S. S., Kramer, C. M., Blumenthal, R. S., & Marvel, F. A. (2022). Medicine 2032: The future of cardiovascular disease prevention with machine

learning and digital health technology. American Journal of Preventive Cardiology, 12, 100379. https://doi.org/10.1016/j.ajpc.2022.100379

[9] Adler, E. D., Voors, A. A., Klein, L., Macheret, F., Braun, O. O., Urey, M. A., Zhu, W., Sama, I., Tadel, M., Campagnari, C., Greenberg, B., & Yagil, A. (2019). Improving risk prediction in heart failure using machine learning. European Journal of Heart Failure, 22(1), 139–147. https://doi.org/10.1002/ejhf.1628

[10] Ali, M. M., Al-Doori, V. S., Mirzah, N., Hemu, A. A., Mahmud, I., Azam, S., Al-Tabatabaie, K. F., Ahmed, K., Bui, F. M., & Moni, M. A. (2023). A machine learning approach for risk factors analysis and survival prediction of Heart Failure patients. Healthcare Analytics, 3, 100182. https://doi.org/10.1016/j.health.2023.100182

[11] Wang, J. (2021). Heart Failure Prediction with Machine Learning: A Comparative Study. Journal of Physics Conference Series, 2031(1), 012068. https://doi.org/10.1088/1742-6596/2031/1/012068

[12] Saqib, M., Perswani, P., Muneem, A., Mumtaz, H., Neha, F., Ali, S., & Tabassum, S. (2024). Machine learning in heart failure diagnosis, prediction, and prognosis: review. Annals of Medicine and Surgery. https://doi.org/10.1097/ms9.0000000000002138

[13] Chicco, Davide, and Giuseppe Jurman. ”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 (2020): 16.

[14] Jain, A. (2021, December 16). Support Vector Machines(S.V.M) — hyperplane and margins. Medium. https://medium.com/@apurvjain37/support-vector-machines-s-v-m-hyperplane-and-margins-ee2f083381b4

[15] Schölkopf, B., & Smola, A. J. (2002). Learning with kernels: Support vector machines, regularization, optimization, and beyond. Retrieved from https://alex.smola.org/papers/2002/SchSmo02b.pdf