

A Machine Learning Approach to Heart Failure Detection

Alban Broze

*abroze, 06510556
CME216, Winter 2023*

Introduction

The objective of this report is to apply machine learning to predict heart failure (HF) based on patient data and evaluate the effectiveness of the model in identifying patients at high risk of heart failure.

Bioengineering is an interdisciplinary field that combines principles of engineering and biology to solve complex problems in healthcare and medicine. Heart failure is a major issue in bioengineering as it affects millions of people worldwide and presents a significant challenge for diagnosis and treatment. The use of machine learning can aid in early detection and improve patient outcomes [1], [2].

In this project, a support vector machine (SVM) classifier is used to predict heart failure. Patients are classified as either high-risk or low-risk for heart failure based on their medical history and other relevant data. This method has been shown to be effective in a wide range of applications, including medical diagnosis and disease prediction.

Using a machine learning approach to heart failure detection can be especially useful in reducing the financial burden associated with medical treatments. Furthermore, it can help identifying the positive cases early on, which is still a very hard task [1] and can help improve the detection accuracy, in a world where a large number of HF patients are still incorrectly diagnosed [2].

Literature Review

In recent years, multiple researchers have focused on using machine learning techniques to predict heart failure for patients. Paper [3] presents an up-to-date overview of the application of machine learning methods in heart failure including diagnosis, classification, readmissions and medication adherence.

Furthermore, [4] presents the state-of-the-art of the machine learning methodologies applied for the assessment of heart failure. SVM is an often used technique for classification tasks in the bioengineering domain. Examples of papers where SVM is used for predicting heart failure include [5] and [6]. Therefore, it seemed relevant to explore it more in detail in the context of this project.

Dataset Description

The dataset used in this study was obtained from Kaggle [7] and consists of medical data from about 300 patients with and without heart failure. The dataset includes 11 features that are relevant for predicting heart failure, as described below in Table 1:

Feature (measure)	Description
Age (years)	Age of the patient
Anaemia (bool)	Indicates whether the patient has low levels of haemoglobin in their blood
CPK (mcg/L)	Level of the CPK enzyme in the blood
Diabetes (bool)	Indicates whether the patient has diabetes
Ejection Fraction (%)	Percentage of blood leaving the heart at each contraction
High Blood Pressure (bool)	Indicates whether the patient has hypertension
Platelets (kiloplatelets/mL)	Platelet count of blood, they are cells in the blood that help with clotting
Serum Creatinine (mg/dL)	Level of creatinine in the blood
Serum Sodium (mEq/L)	Level of sodium in the blood
Sex (bool)	Indicates the patient's gender
Smoking (bool)	Indicates whether the patient is a smoker

Table 1: Feature description

Approach

The approach used in this project to classify patients as at risk of heart failure or not is SVM with an 'RBF' kernel on mean-normalized data. SVM is a machine learning algorithm that separates data points into different classes using a boundary. In this case, the boundary separates patients at risk of heart failure from those who are not.

SVM involves hyperparameters like 'C' and 'gamma'. 'C' controls the trade-off between correctly classifying training examples and maximizing the decision boundary. 'Gamma' defines the influence of a single training example, with high 'gamma' leading to the model overfitting the data. The 'RBF' kernel function transforms the input data into a higher dimensional space, where it can be more easily separated.

To find the optimal combination of hyperparameters, a grid search was implemented for values of 'C' = [0.01, 0.1, 1, 5, 10, 20] and 'gamma' = [1, 0.1, 0.05, 0.01, 0.001, 0.0001,

0.00001]. This tuning process ensures the model will generalize well to unseen data. The best 'C' - 'gamma' pair was found to be 'C' = 10 and 'gamma' = 0.0001.

In order to make sure that the code is correctly implementing SVM, multiple actions have been taken. First, the grid search steps have been printed in order to make sure that the model is considering a broad range of parameter pairs to find the best possible pair of 'C' and 'gamma'. Second, the results have been plotted to assess consistency of the predictions with the actual data. Finally, the data has been split in training and testing sets in random ways to assess model consistency, and no significant result changes were noticed. Eventually, random state 25 was chosen (randomly) to always achieve similar results. This helped in tuning the hyperparameters 'C' and 'gamma' during the grid search.

Results and Interpretation

This section will present an overview of the obtained results of using SVM as a method to detect risk of heart failures for patients. As can be seen from Figure 1, the model does a good job at correctly classifying heart failures cases. Indeed, both the number of negative cases (0) and positive cases (1) for the predictions are close to the actual data.

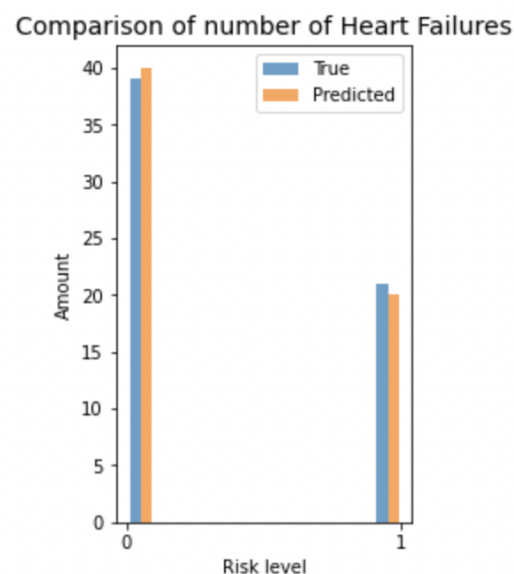


Figure 1: True vs. Predicted Values for Heart Failure

To dive more into the results, we can analyze the classification report. The classification report consists in analyzing the precision, recall and F1-score of our model. "Precision" tells us what proportion of positive identifications was actually correct, "recall" tells us what proportion of actual positives was identified correctly, and the "F1-score" is the harmonic mean of the "precision" and the "recall". The metric is defined as follows:

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall}$$

As can be seen from the classification report in Figure 2, the model scores well on all three metrics, with weighted averages equal to 0.85 for all of them.

	precision	recall	f1-score	support
0	0.88	0.90	0.89	39
1	0.80	0.76	0.78	21
accuracy			0.85	60
macro avg	0.84	0.83	0.83	60
weighted avg	0.85	0.85	0.85	60

Figure 2: Classification Report

To complete this analysis, we can plot the Confusion Matrix. This is a very useful visualization to help understand which parts of the data have been the most misclassified and misunderstood by the model. To Confusion Matrix for this project can be seen in Figure 3 below.

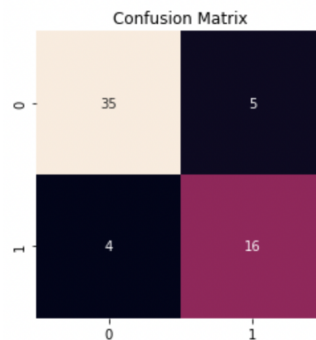


Figure 3: Confusion Matrix

What can be seen from this figure is that 35 actual negative cases were also classified as such by the model, while 5 actual negative cases were misclassified as positive cases. On the other hand, 16 actual positive cases were classified as such by the model, while 4 actual positive cases were misclassified as negative cases. Depending on the application, one could prefer having false positives (FP) rather than false negatives (FN), or vice versa. In our case, if we consider that not identifying an actual positive cases should absolutely avoided, we should try to minimize the number of false negatives (i.e. bottom left cell).

To conclude, we can say that the model is able to characterize the general trend in the data well. The predictions and the scores are satisfying and promising. Nevertheless, it could be beneficial to increase the model's complexity to reduce the number of FP and FN.

Limitations and Next Steps

Even if the above presented results are satisfying and promising, this model still has its limitations and additional work could be done to increase its accuracy. First, limitations will be described and then next steps will be suggested.

The following are limitations of this model and its implementation:

1. **Features:** Only the features provided in the original dataset have been used to build the model. However, other variables could have a role in heart failure cases, like the heartbeat for example. It might have been better to engineer new features to complexify the model. But because feature engineering is an entire domain of research, this hasn't been done to avoid getting lost in an over-fitted model.
2. **Data:** Only about 300 samples were used to build the model. This means that the data hasn't been trained on that many samples, and that the test set wasn't that big either. More data would have been nice. Furthermore, the data has been retrieved from Kaggle, but comes from an unknown source and there is no information about about how the data was gathered and filtered.
3. **Model:** For the sake of simplicity, only one method (SVM) has been tested to build a reliable model. However, other methods might work better. Furthermore, only the 'RBF' kernel was tested for the SVM model, while others might work better. 'RBF' was chosen because it's considered the 'go-to' kernel for SVM [8].

Next steps to improve the model and be able to better predict heart failure cases could include: doing featuring engineering to obtain the best features possible, using more data from reliable sources, and testing out other methods than SVM.

Furthermore, signal-based approaches could be used to detect heart failure. For example, by looking at collected ECG and SCG signals, processing them and using machine learning algorithms to extract data trends, one could potentially better identify heart failure risk [9].

Conclusion

Machine learning-based heart failure detection is promising and has potential to improve accuracy and timeliness of diagnoses. As technology advances and more data becomes available, we can expect even more sophisticated algorithms and models to be developed. Exciting times are ahead for this field, and I look forward to reading about future improvements and breakthroughs in the diagnosis and treatment of heart failure. With continued research and innovation, we can improve outcomes for patients and ultimately save lives.

References

- [1] Fonseca, C. (2006). Diagnosis of heart failure in primary care. *Heart Failure Reviews*, 11, 95-107.
- [2] Fuat, A., Hungin, A. P. S., & Murphy, J. J. (2003). Barriers to accurate diagnosis and effective management of heart failure in primary care: Qualitative study. *BMJ*, 326, 196-201.
- [3] Baksi, A. J., & Pennell, D. J. (2018). Machine learning in heart failure – ready for prime time? *Current Opinion in Cardiology*, 33(2), 190-195.
- [4] Tripoliti, E. E., Papadopoulos, T. G., Karanasiou, G. S., Naka, K. K., & Fotiadis, D. I. (2017). Heart failure: Diagnosis, severity estimation and prediction of adverse events through machine learning techniques. *Computational and Structural Biotechnology Journal*, 15, 26-47.
- [5] Yang, G., Ren, Y., Pan, Q., Ning, G., Gong, S., Cai, G., Zhang, Z., Li, L., & Yan, J. (2010). A heart failure diagnosis model based on support vector machine. In 3rd International Conference on Biomedical Engineering and Informatics.
- [6] Gnaneswar, B., & Ebenezar Jebarani, M. R. (2017). A review on prediction and diagnosis of heart failure. In 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS).
- [7] Andrew M. V. (2020). Heart Failure Clinical Data [Data set]. Kaggle. <https://www.kaggle.com/datasets/andrewmvd/heart-failure-clinical-data>
- [8] Sreenivasa, S. (2020, October 12). Radial Basis Function (RBF) Kernel: The Go-To Kernel. Towards Data Science. <https://towardsdatascience.com/radial-basis-function-rbf-kernel-the-go-to-kernel-acf0d22c798a>
- [9] HeartKinetics. (n.d.). Home. Retrieved March 13, 2023, from <https://heartkinetics.com/>

Import Libraries

```
In [349... import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.svm import SVC
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
from sklearn.model_selection import GridSearchCV
from sklearn.model_selection import train_test_split
```

Import Data

```
In [350... hf = pd.read_csv('heart_failure_data/heart_failure_clinical_records_dataset.csv')
```

Print Features & Target Variable

```
In [351... print("Features: \n", np.array(hf.columns[:-1]), "\n")
print("Target variable: \n", hf.columns[-1])
```

```
Features:
['age' 'anaemia' 'creatinine_phosphokinase' 'diabetes' 'ejection_fraction'
 'high_blood_pressure' 'platelets' 'serum_creatinine' 'serum_sodium' 'sex'
 'smoking' 'time']
```

```
Target variable:
DEATH_EVENT
```

Split and Normalize the Data

```
In [352... # split the data in training and testing set (20%)
hf_train, hf_test = train_test_split(hf, test_size=0.2, random_state=25)

# split training and testing data into X and y
X_train, y_train = hf_train.iloc[:, :-1], hf_train.iloc[:, -1]
X_test, y_test = hf_test.iloc[:, :-1], hf_test.iloc[:, -1]

# normalize data using mean normalization
X_train = (X_train - X_train.mean()) / X_train.std()
X_test = (X_test - X_test.mean()) / X_test.std()
```

Construct Model

```
In [353... model = SVC(kernel='rbf', class_weight='balanced')
```

Prepare the parameters for a GridSearch

```
In [354... param_grid = {'C': [0.01, 0.1, 1, 5, 10, 20],
                 'gamma': [1, 0.1, 0.05, 0.01, 0.001, 0.0001, 0.00001]}
```

```
In [355]: grid = GridSearchCV(model, param_grid, refit = True, verbose=2)
grid.fit(X_train, y_train)
print("\n Best parameters: \n", grid.best_params_)
```

Fitting 5 folds for each of 42 candidates, totalling 210 fits

```
[CV] END .....C=0.01, gamma=1; total time= 0.0s
[CV] END .....C=0.01, gamma=1; total time= 0.0s
[CV] END .....C=0.01, gamma=1; total time= 0.0s
[CV] END .....C=0.01, gamma=1; total time= 0.0s
[CV] END .....C=0.01, gamma=1; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1; total time= 0.0s
[CV] END .....C=0.01, gamma=0.05; total time= 0.0s
[CV] END .....C=0.01, gamma=0.05; total time= 0.0s
[CV] END .....C=0.01, gamma=0.05; total time= 0.0s
[CV] END .....C=0.01, gamma=0.05; total time= 0.0s
[CV] END .....C=0.01, gamma=0.05; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01; total time= 0.0s
[CV] END .....C=0.01, gamma=0.001; total time= 0.0s
[CV] END .....C=0.01, gamma=0.001; total time= 0.0s
[CV] END .....C=0.01, gamma=0.001; total time= 0.0s
[CV] END .....C=0.01, gamma=0.001; total time= 0.0s
[CV] END .....C=0.01, gamma=0.001; total time= 0.0s
[CV] END .....C=0.01, gamma=0.0001; total time= 0.0s
[CV] END .....C=0.01, gamma=0.0001; total time= 0.0s
[CV] END .....C=0.01, gamma=0.0001; total time= 0.0s
[CV] END .....C=0.01, gamma=0.0001; total time= 0.0s
[CV] END .....C=0.01, gamma=0.0001; total time= 0.0s
[CV] END .....C=0.01, gamma=1e-05; total time= 0.0s
[CV] END .....C=0.01, gamma=1e-05; total time= 0.0s
[CV] END .....C=0.01, gamma=1e-05; total time= 0.0s
[CV] END .....C=0.01, gamma=1e-05; total time= 0.0s
[CV] END .....C=0.01, gamma=1e-05; total time= 0.0s
[CV] END .....C=0.1, gamma=1; total time= 0.0s
[CV] END .....C=0.1, gamma=1; total time= 0.0s
[CV] END .....C=0.1, gamma=1; total time= 0.0s
[CV] END .....C=0.1, gamma=1; total time= 0.0s
[CV] END .....C=0.1, gamma=1; total time= 0.0s
[CV] END .....C=0.1, gamma=0.1; total time= 0.0s
[CV] END .....C=0.1, gamma=0.1; total time= 0.0s
[CV] END .....C=0.1, gamma=0.1; total time= 0.0s
[CV] END .....C=0.1, gamma=0.1; total time= 0.0s
[CV] END .....C=0.1, gamma=0.1; total time= 0.0s
[CV] END .....C=0.1, gamma=0.05; total time= 0.0s
[CV] END .....C=0.1, gamma=0.05; total time= 0.0s
[CV] END .....C=0.1, gamma=0.05; total time= 0.0s
[CV] END .....C=0.1, gamma=0.05; total time= 0.0s
[CV] END .....C=0.1, gamma=0.05; total time= 0.0s
[CV] END .....C=0.1, gamma=0.01; total time= 0.0s
[CV] END .....C=0.1, gamma=0.01; total time= 0.0s
[CV] END .....C=0.1, gamma=0.01; total time= 0.0s
[CV] END .....C=0.1, gamma=0.01; total time= 0.0s
[CV] END .....C=0.1, gamma=0.01; total time= 0.0s
[CV] END .....C=0.1, gamma=0.001; total time= 0.0s
[CV] END .....C=0.1, gamma=0.001; total time= 0.0s
[CV] END .....C=0.1, gamma=0.001; total time= 0.0s
```


[illegible]

[illegible]

```
[CV] END .....C=20, gamma=0.01; total time= 0.0s
[CV] END .....C=20, gamma=0.01; total time= 0.0s
[CV] END .....C=20, gamma=0.01; total time= 0.0s
[CV] END .....C=20, gamma=0.01; total time= 0.0s
[CV] END .....C=20, gamma=0.01; total time= 0.0s
[CV] END .....C=20, gamma=0.001; total time= 0.0s
[CV] END .....C=20, gamma=0.001; total time= 0.0s
[CV] END .....C=20, gamma=0.001; total time= 0.0s
[CV] END .....C=20, gamma=0.001; total time= 0.0s
[CV] END .....C=20, gamma=0.001; total time= 0.0s
[CV] END .....C=20, gamma=0.001; total time= 0.0s
[CV] END .....C=20, gamma=0.0001; total time= 0.0s
[CV] END .....C=20, gamma=0.0001; total time= 0.0s
[CV] END .....C=20, gamma=0.0001; total time= 0.0s
[CV] END .....C=20, gamma=0.0001; total time= 0.0s
[CV] END .....C=20, gamma=0.0001; total time= 0.0s
[CV] END .....C=20, gamma=0.0001; total time= 0.0s
[CV] END .....C=20, gamma=1e-05; total time= 0.0s
[CV] END .....C=20, gamma=1e-05; total time= 0.0s
[CV] END .....C=20, gamma=1e-05; total time= 0.0s
[CV] END .....C=20, gamma=1e-05; total time= 0.0s
[CV] END .....C=20, gamma=1e-05; total time= 0.0s
```

```
Best parameters:
{'C': 10, 'gamma': 0.0001}
```

Make Predictions

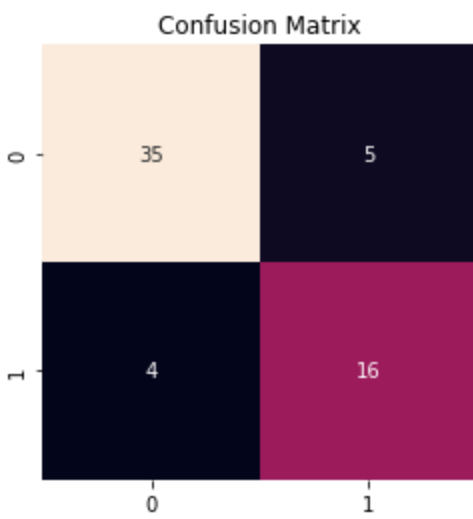
```
In [356... y_pred = grid.predict(X_test)
```

Results

```
In [357... print(classification_report(y_test, y_pred))
conf_mat = confusion_matrix(y_test, y_pred)
sns.heatmap(conf_mat.T, square=True, annot=True, fmt='d', cbar=False).set(title='Confusi
```

	precision	recall	f1-score	support
0	0.88	0.90	0.89	39
1	0.80	0.76	0.78	21
accuracy			0.85	60
macro avg	0.84	0.83	0.83	60
weighted avg	0.85	0.85	0.85	60

```
Out[357]: [Text(0.5, 1.0, 'Confusion Matrix')]
```



```
In [358... plt.figure(figsize=(3, 5))
plt.title("Comparison of number of Heart Failures", size=(14))
plt.ylabel("Amount")
plt.xlabel("Risk level")
plt.hist([y_test, y_pred], alpha=0.7, label=["True", "Predicted"])
plt.xticks([0, 1])
plt.legend()
plt.show()
```

Comparison of number of Heart Failures

