
Heart Disease Classifier

— Justin Short —

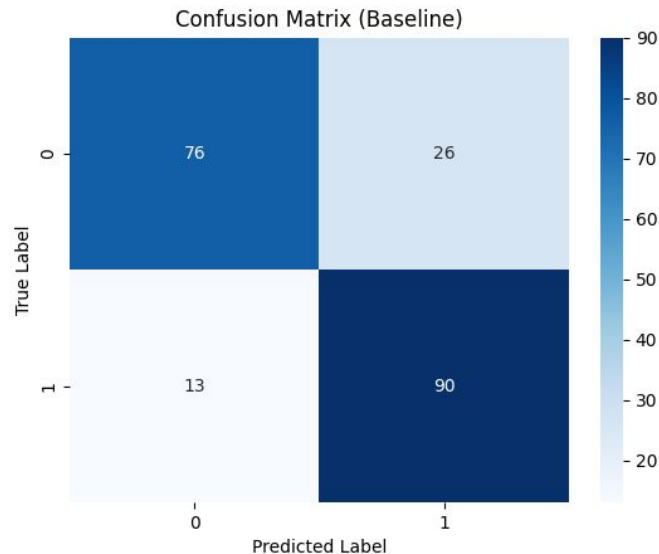
Overview

- SVM Model from scikit-learn
- Dataset from kaggle
 - Heart Disease Dataset
 - <https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset>
 - 13 features including: sex, age, resting heart rate, cholesterol, etc.
 - Target: 1 = "Heart Disease", 0 = "No Heart Disease"
- Model Performance
- Visualizations
- Github
 - https://github.com/jnshort/heart_disease_classifier

Baseline Model Performance

	Precision	Recall	F1-Score	Support
0	0.85	0.75	0.8	102
1	0.78	0.87	0.82	103

Overall Accuracy: 0.81



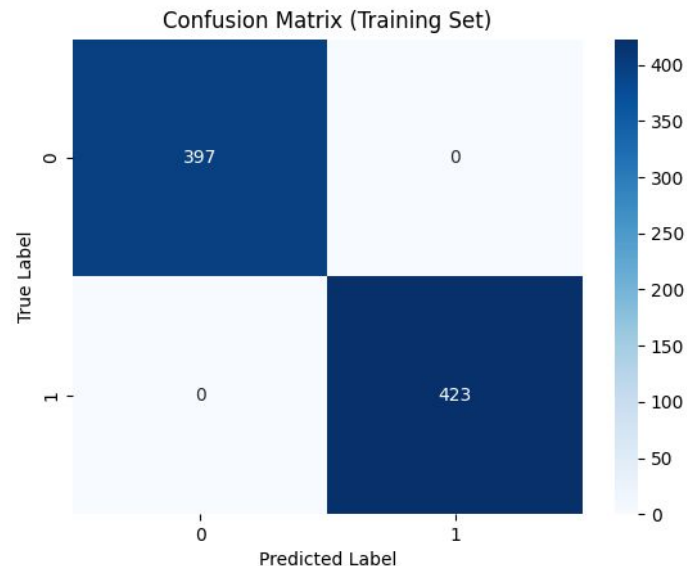
Model Improvements

- Cross validation to extract more out of the limited number of instances in the dataset
- Hyperparameter tuning
 - Using gridsearch
 - Optimal Parameters: C, gamma, kernel
 - C = 0.1
 - Gamma = 1
 - Kernel = polynomial
- (Note: these steps were performed together)

Tuned Model Performance (Train)

	Precision	Recall	F1-Score	Support
0	1.0	1.0	1.0	397
1	1.0	1.0	1.0	423

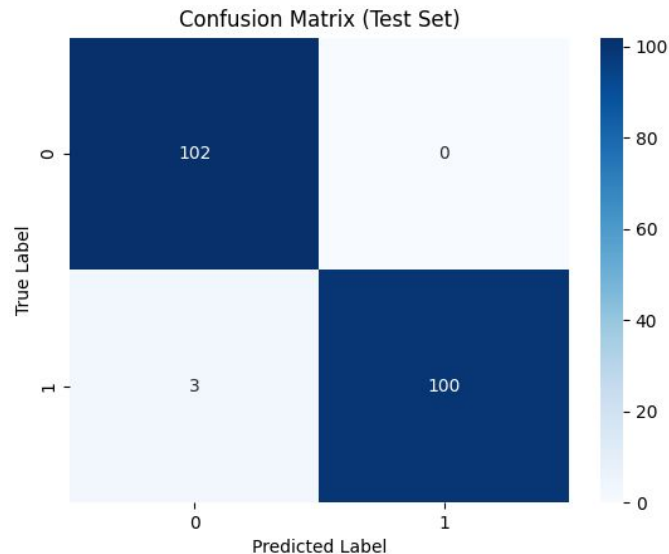
Overall Accuracy: 1.00



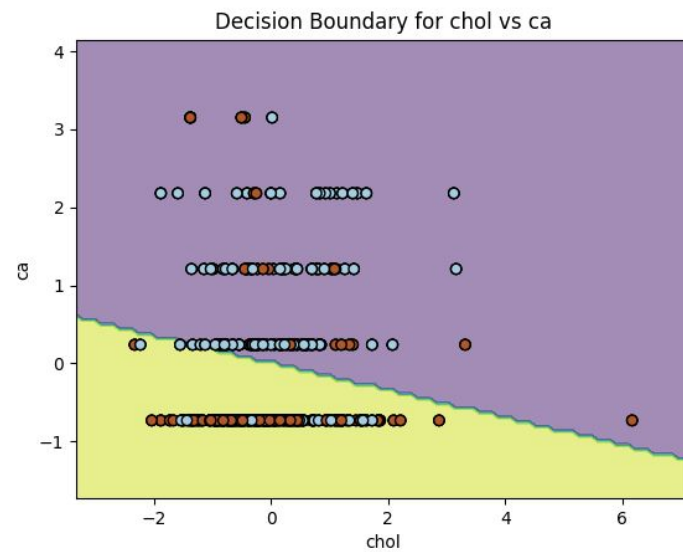
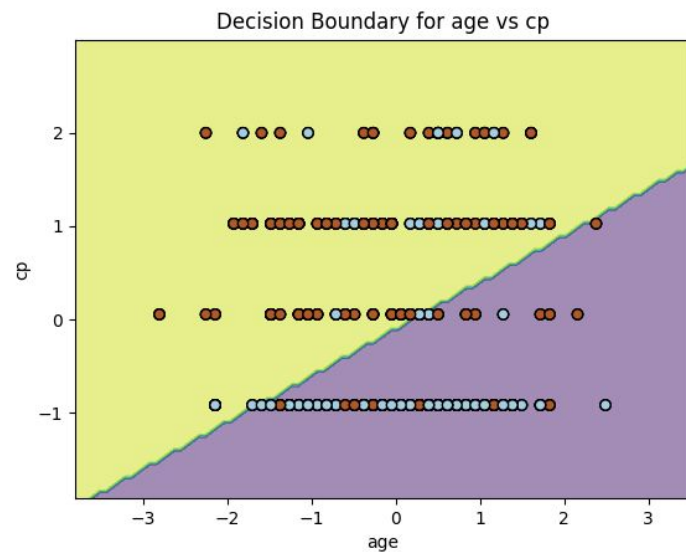
Tuned Model Performance (Test)

	Precision	Recall	F1-Score	Support
0	0.97	1.0	0.99	102
1	1.0	0.97	0.99	103

Overall Accuracy: 0.99



Decision Boundaries



Thank You!