

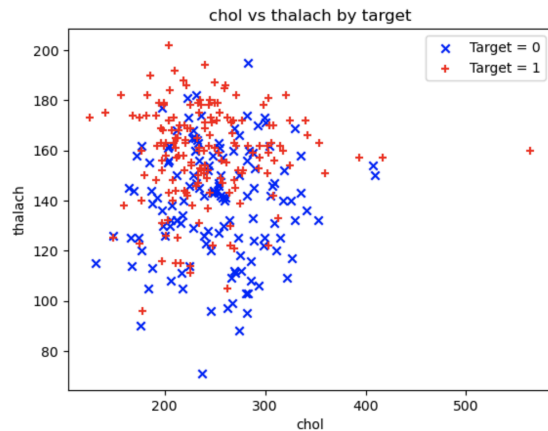
CSCI183_HW3_Report

Part I: Heart Attack Prediction

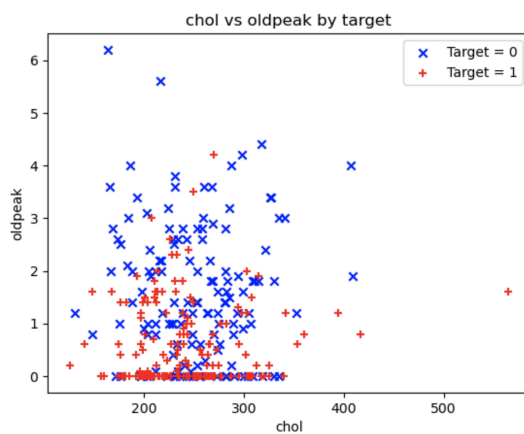
Most relevant numerical features selected in pairs/group, and the scatter plot of each:

And more other plots(including cat-target) included in code.

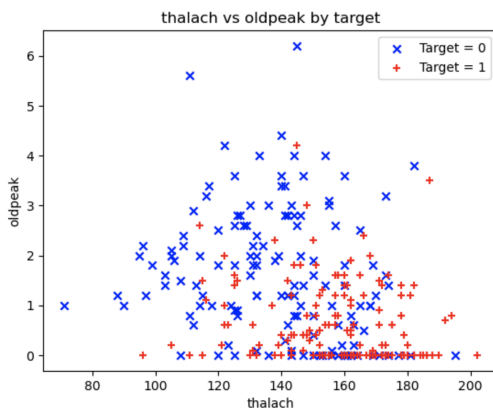
(chol, thalach)



(chol, oldpeak)



(thalach, oldpeak)



Since these three are the most relevant, I also did a logistic regression with three features, but no scatter plot for this. # (chol, thalach, oldpeak)

Observation table in panda_dataframe:

	Features	Precision	Recall	Accuracy	F1 Score	Confusion Matrix
0	chol, thalach	0.6833333333	0.8723404255	0.7252747253	0.7663551402	[[25, 19], [6, 41]]
1	chol, oldpeak	0.6000000000	0.8297872340	0.6263736264	0.6964285714	[[18, 26], [8, 39]]
2	thalach, oldpeak	0.6964285714	0.8297872340	0.7252747253	0.7572815534	[[27, 17], [8, 39]]
3	chol, thalach, oldpeak	0.6964285714	0.8297872340	0.7252747253	0.7572815534	[[27, 17], [8, 39]]