

APPENDIX

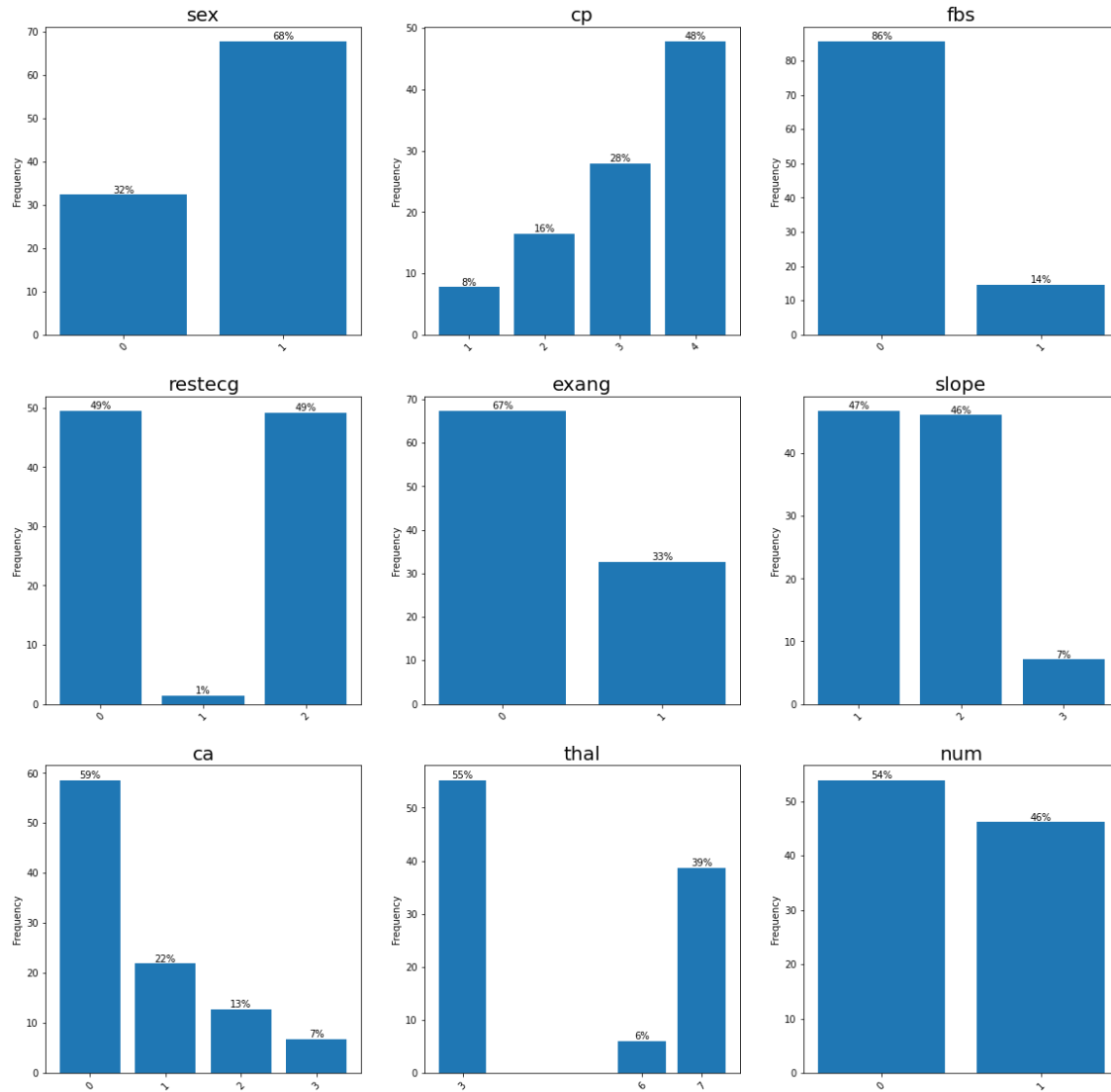


Figure 1. Distribution of Discrete features

	age	trestbps	chol	thalach	oldpeak
count	297.000000	297.000000	297.000000	297.000000	297.000000
mean	54.542088	131.693603	247.350168	149.599327	1.055556
std	9.049736	17.762806	51.997583	22.941562	1.166123
min	29.000000	94.000000	126.000000	71.000000	0.000000
25%	48.000000	120.000000	211.000000	133.000000	0.000000
50%	56.000000	130.000000	243.000000	153.000000	0.800000
75%	61.000000	140.000000	276.000000	166.000000	1.600000
max	77.000000	200.000000	564.000000	202.000000	6.200000

Figure 2. Summary Statistics of Numeric features

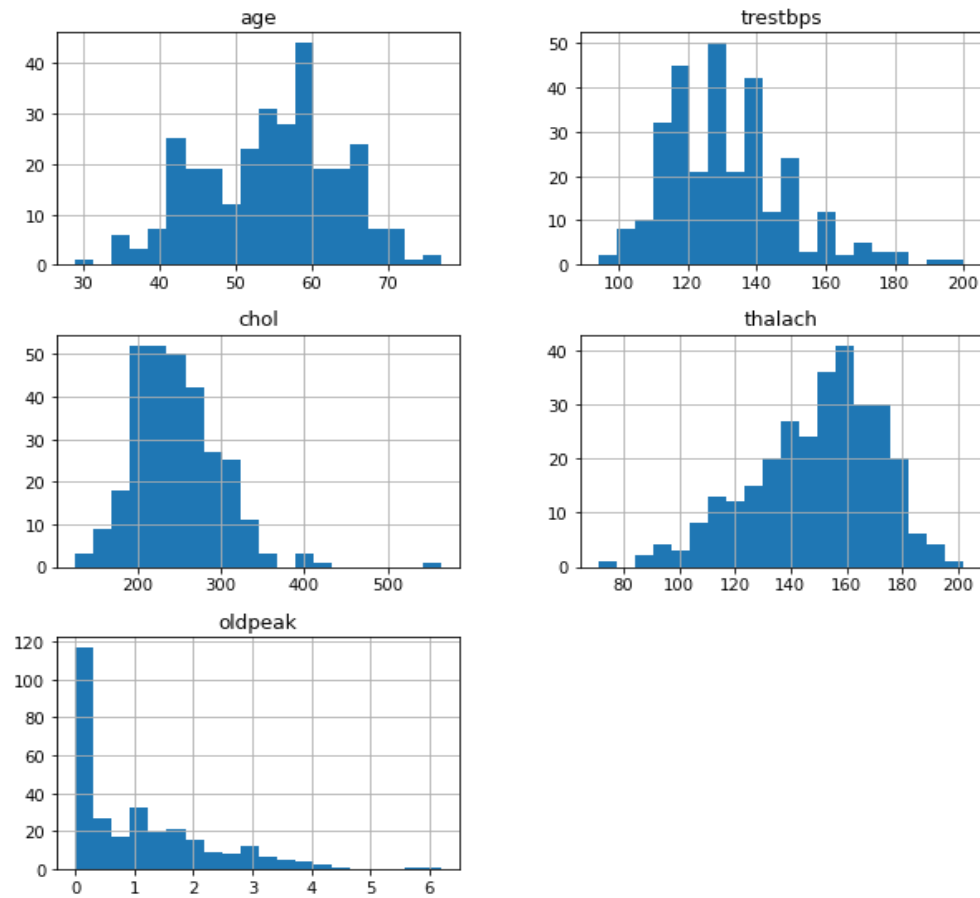


Figure 3. Distribution of Continuous features

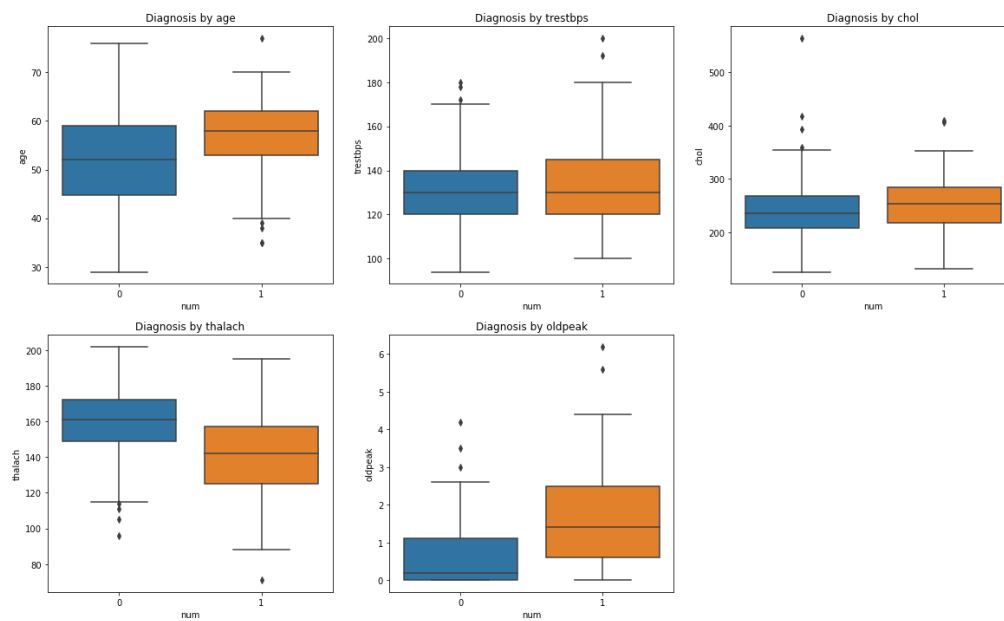


Figure 4. Distribution of label (heart disease diagnosis) by the numeric features

K Nearest Neighbors (KNN) Classifier

	average_score	accuracy	precision	recall	f1	weight	k
0	0.55	0.58	0.54	0.54	0.54	distance	1
1	0.55	0.58	0.54	0.54	0.54	uniform	1
2	0.65	0.68	0.65	0.63	0.64	distance	3
3	0.65	0.68	0.65	0.63	0.64	uniform	3
4	0.66	0.7	0.69	0.61	0.65	distance	5
5	0.67	0.7	0.68	0.63	0.66	uniform	5
6	0.63	0.67	0.65	0.59	0.62	distance	15
7	0.59	0.63	0.61	0.54	0.57	uniform	15
8	0.61	0.66	0.65	0.54	0.59	distance	25
9	0.55	0.6	0.57	0.49	0.53	uniform	25
10	0.56	0.64	0.68	0.41	0.52	distance	50
11	0.56	0.66	0.78	0.34	0.47	uniform	50

Figure 5. Performance Result of KNN Classifier with different combinations of Hyperparameters

Support Vector Machine (SVM) Classifier

	average_score	accuracy	precision	recall	f1	kernel	reg_parameter
0	0.88	0.89	0.92	0.83	0.87	linear	0.025
1	0.89	0.9	0.94	0.83	0.88	linear	0.05
2	0.85	0.87	0.89	0.8	0.85	linear	0.5
3	0.85	0.87	0.89	0.8	0.85	linear	1
4	0.83	0.84	0.89	0.76	0.82	linear	5
5	0.14	0.54	0.0	0.0	0.0	poly	0.025
6	0.14	0.54	0.0	0.0	0.0	poly	0.05
7	0.56	0.66	0.78	0.34	0.47	poly	0.5
8	0.62	0.69	0.76	0.46	0.58	poly	1
9	0.68	0.73	0.79	0.56	0.66	poly	5
10	0.14	0.54	0.0	0.0	0.0	rbf	0.025
11	0.14	0.54	0.0	0.0	0.0	rbf	0.05
12	0.54	0.64	0.76	0.32	0.45	rbf	0.5
13	0.59	0.67	0.76	0.39	0.52	rbf	1
14	0.66	0.71	0.78	0.51	0.62	rbf	5

Figure 6. Performance Result of SVM Classifier with different combinations of Hyperparameters

Decision Tree Classifier

	average_score	accuracy	precision	recall	f1	tree_max_depth
0	0.73	0.76	0.76	0.68	0.72	5.0
1	0.66	0.68	0.64	0.66	0.65	10.0
2	0.66	0.67	0.62	0.68	0.65	15.0
3	0.70	0.71	0.67	0.71	0.69	20.0
4	0.70	0.72	0.69	0.71	0.70	25.0
5	0.70	0.71	0.67	0.73	0.70	30.0

Figure 7. Performance Result of Decision tree with different combinations of Hyperparameters

Random Forest Classifier

	average_score	accuracy	precision	recall	f1	tree_no	max_depth	max_features
0	0.77	0.80	0.83	0.71	0.76	10.0	5.0	1.0
1	0.80	0.82	0.86	0.73	0.79	10.0	5.0	2.0
2	0.84	0.86	0.87	0.80	0.84	10.0	10.0	1.0
3	0.86	0.88	0.92	0.80	0.86	10.0	10.0	2.0
4	0.77	0.80	0.83	0.71	0.76	10.0	15.0	1.0
5	0.81	0.83	0.88	0.73	0.80	10.0	15.0	2.0
6	0.88	0.89	0.94	0.80	0.87	100.0	5.0	1.0
7	0.85	0.87	0.89	0.80	0.85	100.0	5.0	2.0
8	0.88	0.89	0.92	0.83	0.87	100.0	10.0	1.0
9	0.86	0.88	0.89	0.83	0.86	100.0	10.0	2.0
10	0.86	0.87	0.87	0.83	0.85	100.0	15.0	1.0
11	0.86	0.88	0.89	0.83	0.86	100.0	15.0	2.0
12	0.86	0.88	0.94	0.78	0.85	200.0	5.0	1.0
13	0.88	0.89	0.92	0.83	0.87	200.0	5.0	2.0
14	0.86	0.88	0.89	0.83	0.86	200.0	10.0	1.0
15	0.84	0.86	0.87	0.80	0.84	200.0	10.0	2.0
16	0.84	0.86	0.87	0.80	0.84	200.0	15.0	1.0
17	0.84	0.86	0.89	0.78	0.83	200.0	15.0	2.0
18	0.86	0.88	0.92	0.80	0.86	500.0	5.0	1.0
19	0.86	0.88	0.89	0.83	0.86	500.0	5.0	2.0
20	0.86	0.87	0.87	0.83	0.85	500.0	10.0	1.0
21	0.85	0.87	0.89	0.80	0.85	500.0	10.0	2.0
22	0.86	0.87	0.87	0.83	0.85	500.0	15.0	1.0
23	0.84	0.86	0.87	0.80	0.84	500.0	15.0	2.0

Figure 8. Performance Result of Random Forest with different combinations of Hyperparameters

AdaBoost Classifier

	average_score	accuracy	precision	recall	f1	max_estimators	learning_rate
0	0.65	0.69	0.69	0.59	0.63	20.0	0.01
1	0.86	0.88	0.89	0.83	0.86	20.0	0.10
2	0.84	0.86	0.83	0.85	0.84	20.0	1.00
3	0.65	0.69	0.69	0.59	0.63	50.0	0.01
4	0.86	0.87	0.85	0.85	0.85	50.0	0.10
5	0.86	0.87	0.85	0.85	0.85	50.0	1.00
6	0.86	0.88	0.89	0.83	0.86	100.0	0.01
7	0.87	0.88	0.88	0.85	0.86	100.0	0.10
8	0.81	0.82	0.79	0.83	0.81	100.0	1.00

Figure 9. Performance Result of AdaBoost with different combinations of Hyperparameters

Neural Network

	average_score	accuracy	precision	recall	f1	hidden_layers	activation	learning_rate	alpha
0	0.82	0.82	0.78	0.85	0.81	50	relu	0.001	0.1
1	0.76	0.78	0.76	0.76	0.76	50	relu	0.001	1
2	0.72	0.77	0.95	0.51	0.67	50	relu	0.01	0.1
3	0.87	0.88	0.88	0.85	0.86	50	relu	0.01	1
4	0.14	0.54	0.0	0.0	0.0	50	relu	0.1	0.1
5	0.14	0.54	0.0	0.0	0.0	50	relu	0.1	1
6	0.69	0.71	0.68	0.68	0.68	50	tanh	0.001	0.1
7	0.68	0.71	0.7	0.63	0.67	50	tanh	0.001	1
8	0.14	0.54	0.0	0.0	0.0	50	tanh	0.01	0.1
9	0.14	0.54	0.0	0.0	0.0	50	tanh	0.01	1
10	0.64	0.46	0.46	1.0	0.63	50	tanh	0.1	0.1
11	0.5	0.47	0.44	0.59	0.5	50	tanh	0.1	1
12	0.78	0.77	0.69	0.88	0.77	100	relu	0.001	0.1
13	0.66	0.72	0.94	0.41	0.58	100	relu	0.001	1
14	0.74	0.66	0.57	1.0	0.73	100	relu	0.01	0.1
15	0.64	0.7	1.0	0.34	0.51	100	relu	0.01	1
16	0.14	0.54	0.0	0.0	0.0	100	relu	0.1	0.1
17	0.67	0.52	0.49	1.0	0.66	100	relu	0.1	1
18	0.5	0.6	0.61	0.34	0.44	100	tanh	0.001	0.1
19	0.48	0.6	0.63	0.29	0.4	100	tanh	0.001	1
20	0.52	0.63	0.83	0.24	0.38	100	tanh	0.01	0.1
21	0.72	0.74	0.71	0.73	0.72	100	tanh	0.01	1
22	0.64	0.46	0.46	1.0	0.63	100	tanh	0.1	0.1
23	0.14	0.54	0.0	0.0	0.0	100	tanh	0.1	1

Figure 10. Performance Result of Neural Network with different combinations of Hyperparameters

Logistic Regression

	average_score	accuracy	precision	recall	f1	penalty
0	0.87	0.88	0.88	0.85	0.86	l1
1	0.87	0.88	0.88	0.85	0.86	l2
2	0.89	0.9	0.94	0.83	0.88	elasticnet

Figure 11. Performance Result of Logistic Regression with different combinations of Hyperparameters