

Predicting Heart Disease with Other Medical Conditions

Ian Crane

McNulty and Associates

Dataset: NHIS 2013 Sample Adult

- 723 fields – most are health / stress related questions
- ~34,000 Respondents
- ~33,000 Respondents who answered what we needed.

Only include information known to patient's insurance company.

- “Have you ever been you told you have....”
- Sex
- Age
- Ethnicity

The Final Field Set

- Had Hypertension?
- Had high cholesterol (12 mos)?
- Had Angina Pectoris?
- Had Heart Attack?
- Had Stroke?
- Had COPD?
- Had Diabetes?
- Had Chronic Bronchitis?
- Had Weak / Failing Kidneys?
- Age
- Sex
- Ethnicity

The Best Classifiers

Model Name	Accuracy	Precision(HD)	Recall(HD)
LogrithmicRegression	96%	70%	39%
GaussianNB	90%	32%	69%
SVC	96%	66%	46%

The Best Classifiers

Model Name	Accuracy	Precision(HD)	Recall(HD)
Mode(Log, GNB, SVC)	96%	0.68	48%
LogrithmicRegression	96%	70%	39%
GaussianNB	90%	32%	69%
SVC	96%	66%	46%

The Best Classifiers

Model Name	Accuracy	Precision(HD)	Recall(HD)
Mode(Log, GNB, SVC)	96%	0.68	48%
LogrithmicRegression	96%	70%	39%
GaussianNB	90%	32%	69%
SVC	96%	66%	46%

What does this mean?