**Description of the Variables**

|  |  |  |  |
| --- | --- | --- | --- |
| Variable: | Description: | Role | Level |
| age | age | Input | Numeric |
| sex | sex | Input | Binary |
| cp | chest pain type (4 values) | Input | Numeric |
| trestbps | resting blood pressure | Input | Numeric |
| chol | serum cholesterol in mg/dl | Input | Numeric |
| fbs | fasting blood sugar > 120 mg/dl | Input | Binary |
| restecg | resting electrocardiographic results (values 0,1,2) | Input | Numeric |
| thalach | maximum heart rate achieved | Input | Numeric |
| exang | exercise induced angina | Input | Binary |
| oldpeak | oldpeak = ST depression induced by exercise relative to rest | Input | Numeric |
| slope | the slope of the peak exercise ST segment | Input | Numeric |
| ca | number of major vessels (0-3) colored by flourosopy | Input | Numeric |
| thal | thal: 3 = normal; 6 = fixed defect; 7 = reversable defect | Input | Categorical |

Error is 8times larger in anamoly than original data

Data of unknown origin has same error rate as normal than it fits patterns as normal data

One big error in normal data can be anamoly inside normal data and it is perfectly possible

There may be some values which are falsely classified in models