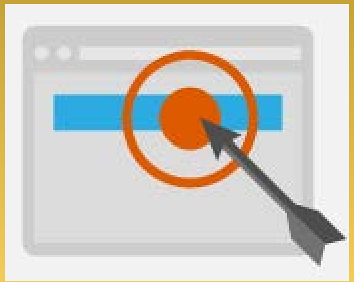
ALGORITHM ANOMALY DETECTION



Algorithm Anomaly detection

In this session

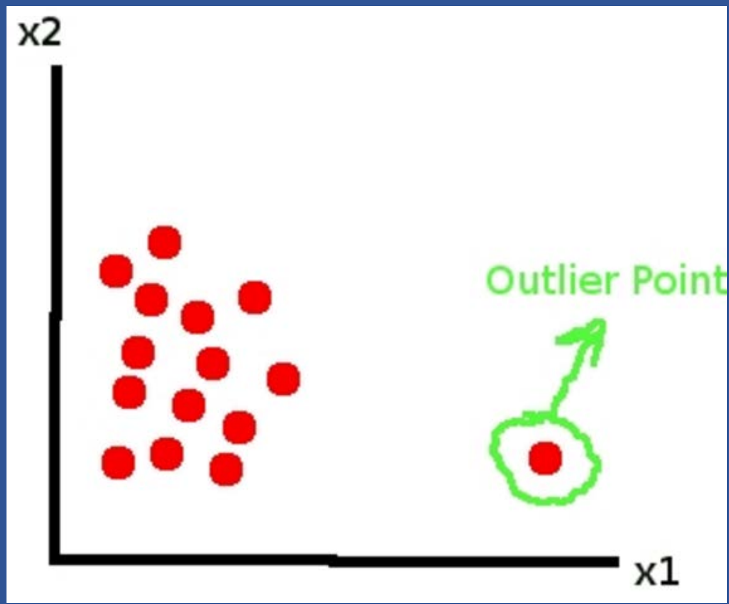
- Anomaly Detection
- One-Class SVM Algorithm
- PCA-Based Algorithm
- Data set
- Data attribute
- Experiment Steps

Algorithm Anomaly detection

Anomaly Detection

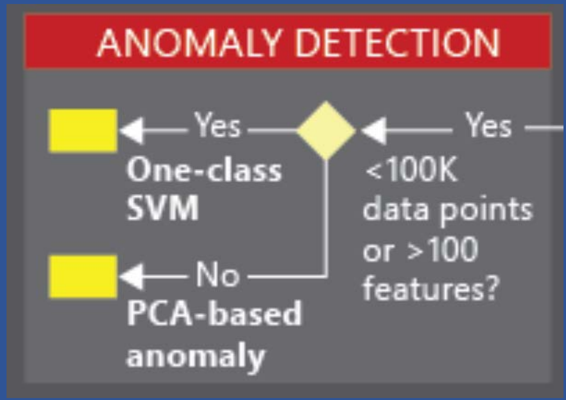
Anomaly Detection

- Credit card fraud, transaction, medical, text etc.
- Also referred to as outliers, novelties, noise, deviations and exceptions
- The data consists of 'normal' applications and 'risky' applications
- Risky transactions = anomalous



Algorithm Anomaly detection

One-Class SVM

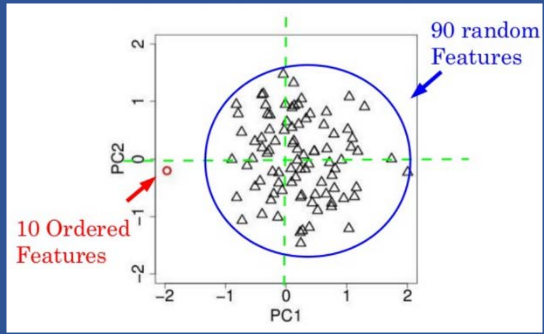


One-Class SVM

- SVM = Support Vector Model
- Supervised learning models
- Analyze data and recognize patterns
- Have a lot of "normal" data and not many cases of the anomalies
- Use with Train Anomaly Detection Model
- The train data set contain all or mostly normal cases.

Algorithm Anomaly detection

PCA-Based



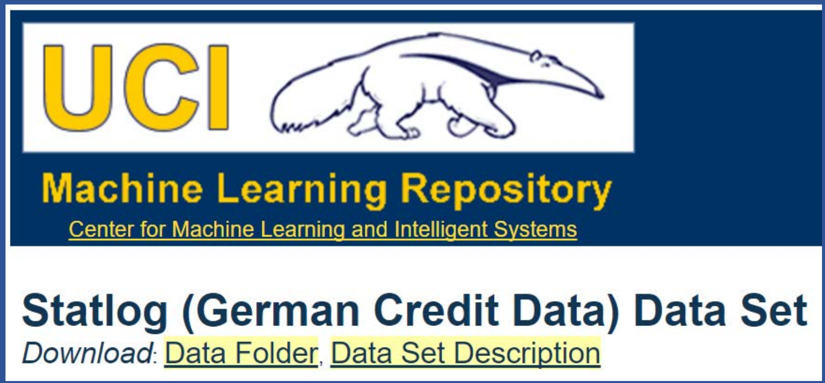
PCA-Based Anomaly Detection module

- Principal Component Analysis (PCA)
- Use when easy to obtain training data from one class
- One class = acceptable transactions
- Use when difficult to obtain sufficient samples of the targeted anomalies
- Detect fraudulent transaction
- You might not have enough examples of fraud to train the mode
- But have many examples of good transactions

Algorithm Anomaly detection

Data set

https://archive.ics.uci.edu/ml/datasets/Statlog+(German+Credit+Data)



Quelle:
Professor Dr. Hans Hofmann
Institut für Statistik und "Okonometrie
Universit "in Hamburg
FB Wirtschaftswissenschaften
Von-Melle-Park 5
2000 Hamburg 13

German credit dataset

- Credit card application
- 1000 instances (rows)
- Attributes = 20 (7 numerical, 13 categorical)
- Label 1 = normal, 2 = risky

Algorithm Anomaly detection

Data attribute

```
A11 6 A34 A43 1169 A65 A75 4 A93 A101 4 A121 67 A143 A152 2 A173 1 A192 A201 1
A12 48 A32 A43 5951 A61 A73 2 A92 A101 2 A121 22 A143 A152 1 A173 1 A191 A201 2
A14 12 A34 A46 2096 A61 A74 2 A93 A101 3 A121 49 A143 A152 1 A172 2 A191 A201 1
A11 42 A32 A42 7882 A61 A74 2 A93 A103 4 A122 45 A143 A153 1 A173 2 A191 A201 1
A11 24 A33 A40 4870 A61 A73 3 A93 A101 4 A124 53 A143 A153 2 A173 2 A191 A201 2
A14 36 A32 A46 9055 A65 A73 2 A93 A101 4 A124 35 A143 A153 1 A172 2 A192 A201 1
A14 24 A32 A42 2835 A63 A75 3 A93 A101 4 A122 53 A143 A152 1 A173 1 A191 A201 1
```

Attribute: Account status, month, credit history, propose, amount, saving, employ since, installment rate, sex …

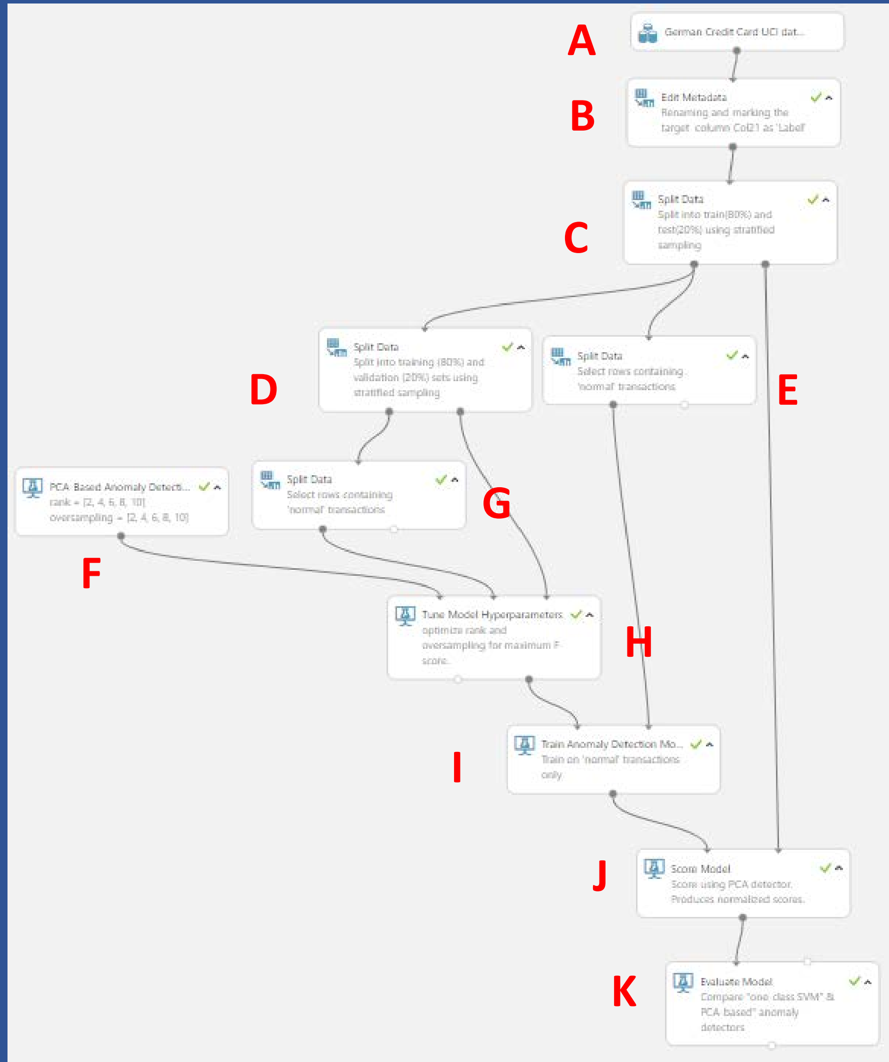
```
Attribute 9:  (qualitative)
          Personal status and sex
          A91 : male   : divorced/separated
          A92 : female : divorced/separated/married
            A93 : male   : single
          A94 : male   : married/widowed
          A95 : female : single
```

Algorithm Anomaly detection

Experiment Steps

Experiment steps

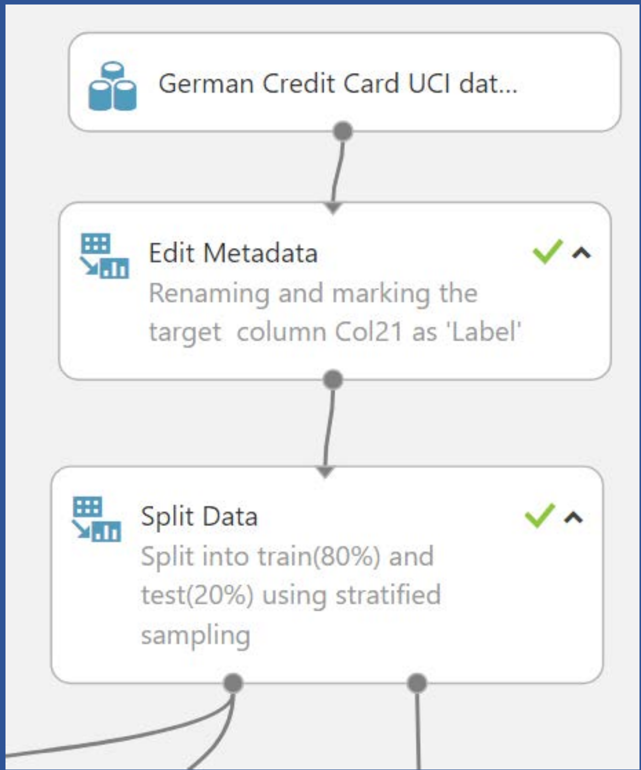
1. Import data set
2. Edit metadata
3. Split data for training
4. Split data for Score
5. Add PCA Base method
6. Add Tune Model Hyper parameters
7. Add Train Anomaly Detection Model
8. Add Score model
9. Add Evaluate Model

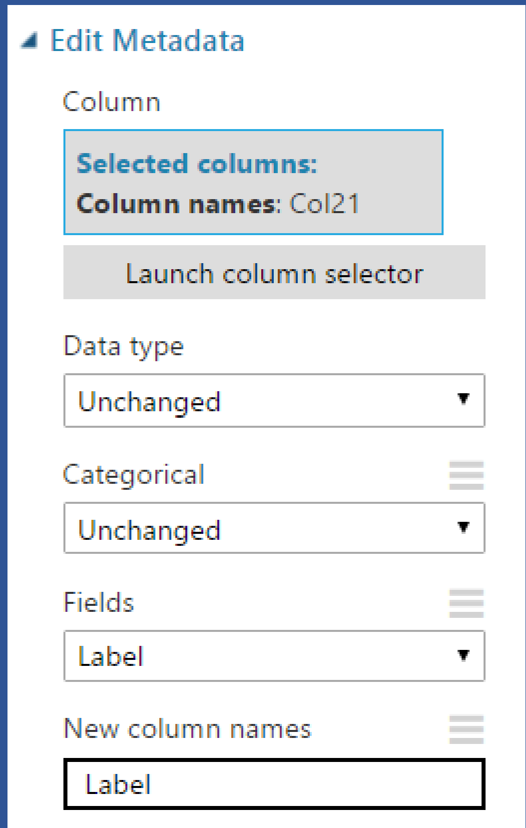


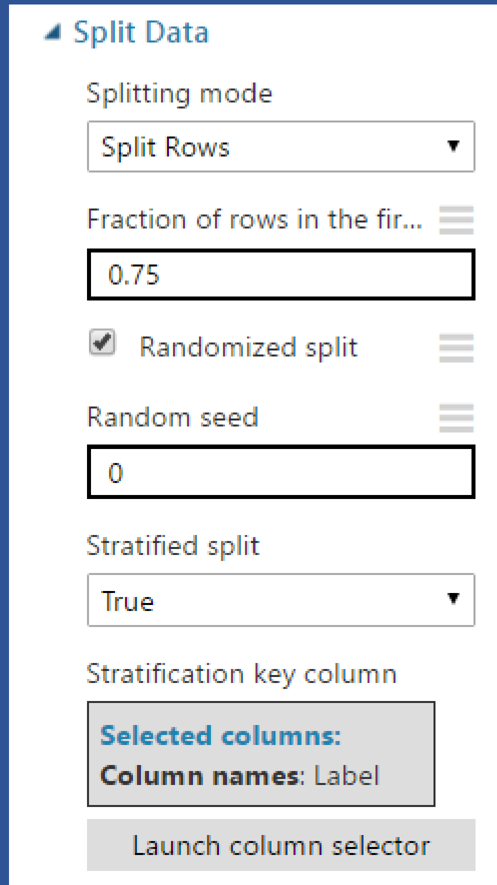
Algorithm Anomaly detection

Experiment Steps

A. Import data set
B. Edit metadata
C. Split data for training





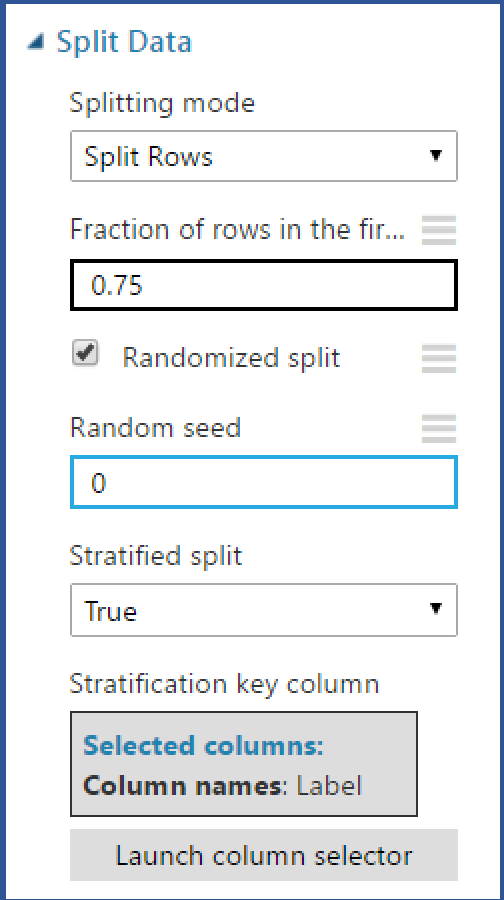


Algorithm Anomaly detection

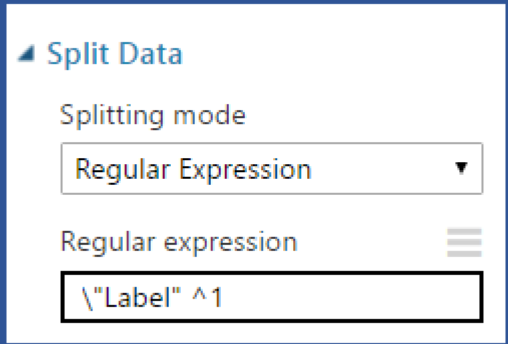
Experiment Steps

Add 4 Split data models and PCA Based

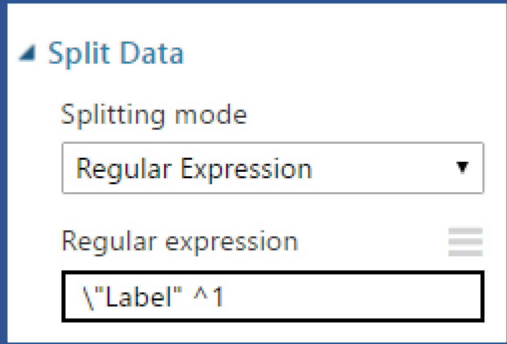
D



E



F



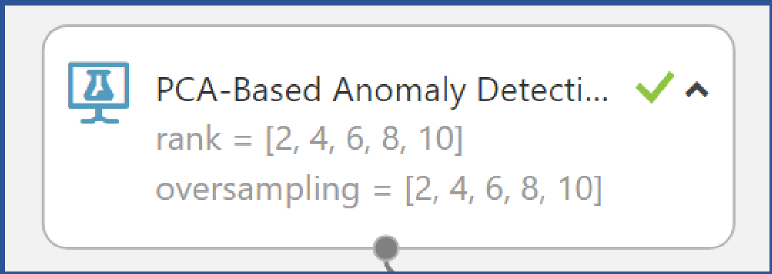
Algorithm Anomaly detection

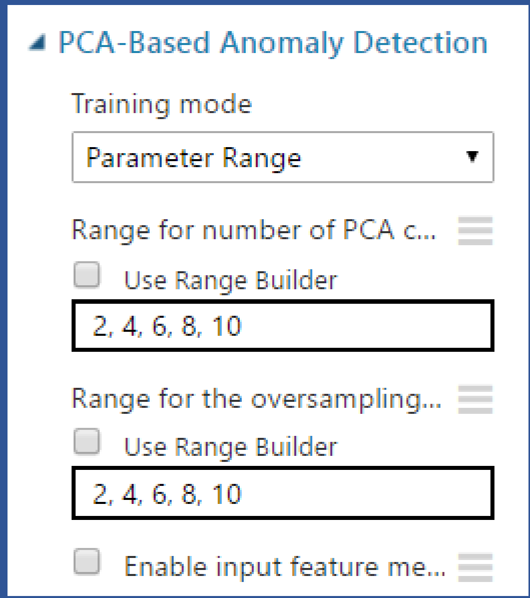
Experiment Steps

Training mode

- Single Parameter: If you know how you want to configure the model, you can provide a specific set of values as arguments. You might have learned these values by experimentation or received them as guidance.
- Parameter Range: If you are not sure of the best parameters, you can find the optimal parameters by specifying multiple values and using a parameter sweep to find the optimal configuration.

G



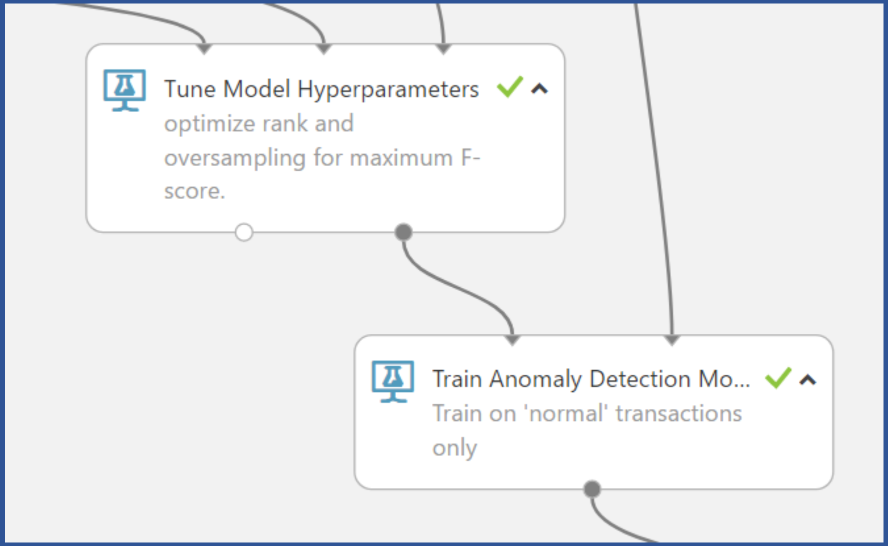


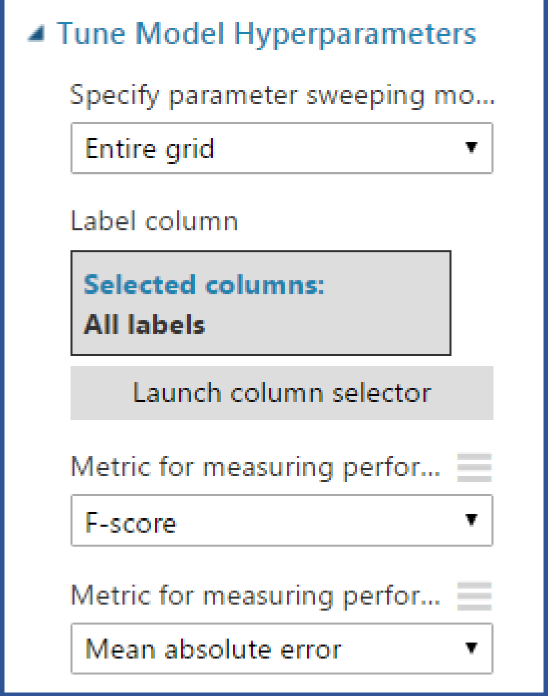
Algorithm Anomaly detection

Experiment Steps

H. Add Tune Model Hyperparameters

I. Add Train Anomaly Detection Model



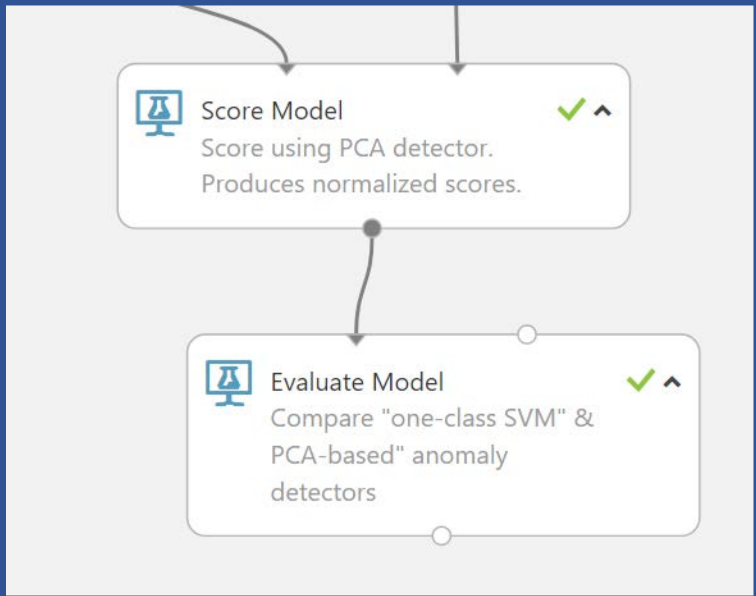


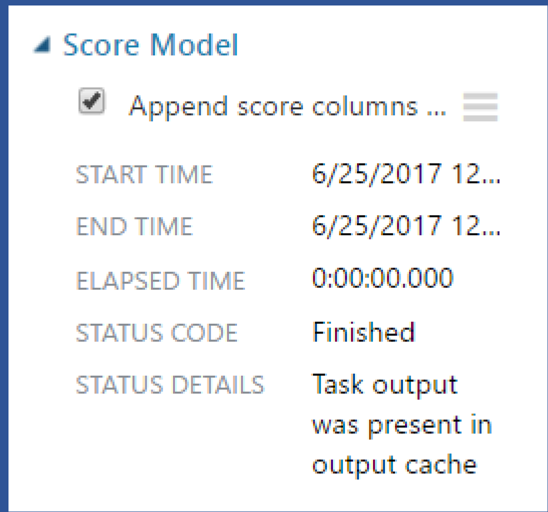
Algorithm Anomaly detection

Experiment Steps

J. Add Score Model

K. Add Evaluate Model



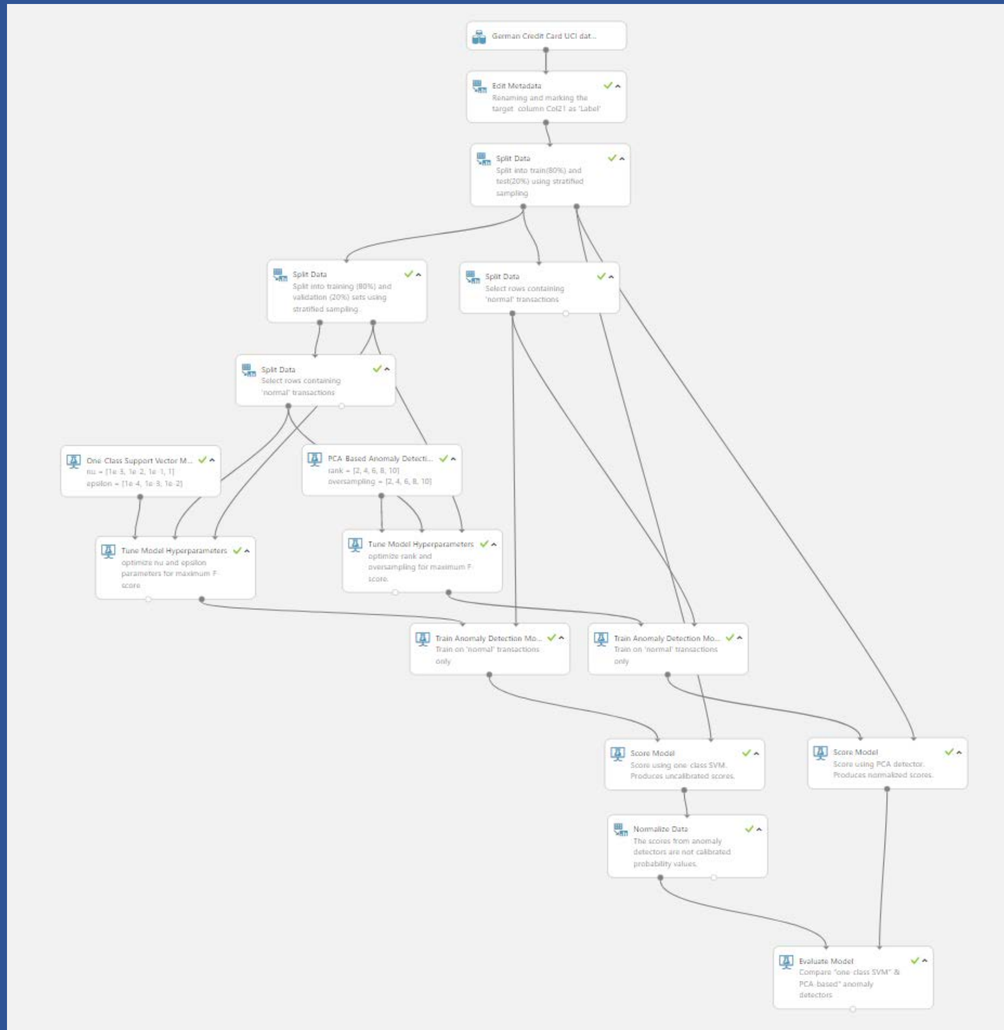


Algorithm Anomaly detection

Compare two anomaly algorithm

Compare One-Class Support Vector Machine with PCA-Based Anomaly Detection

https://gallery.cortanaintelligence.com/Experiment/Anomaly-compare



Algorithm Anomaly detection

More Information

PCA-Based Anomaly Detection

https://msdn.microsoft.com/en-us/library/azure/dn913102.aspx

This Experiment

https://gallery.cortanaintelligence.com/Experiment/Anomaly-Detection-9

Anomaly compare

https://gallery.cortanaintelligence.com/Experiment/Anomaly-compare