

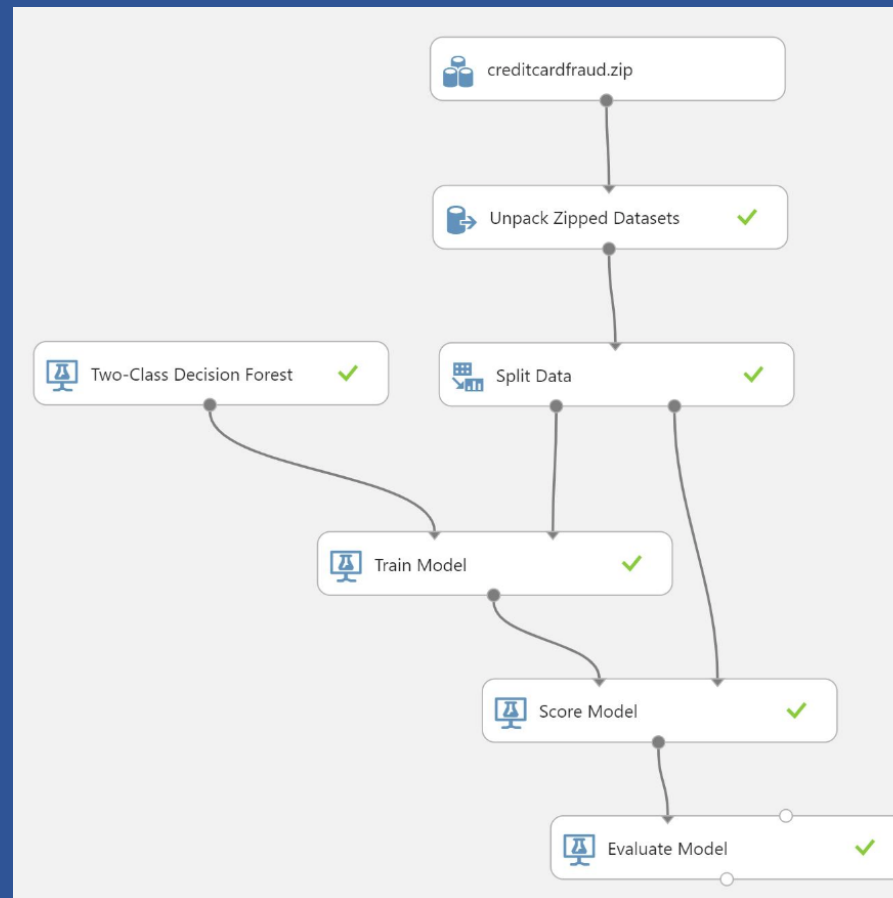
Credit Card in Azure ML

In this session

- Question and Data
- Dataset description
- Create Data Set
- Place dataset
- Drop missing data rows
- Split Data
- Train, Score, Evaluate
- Understand Confusion matrix

The finished model

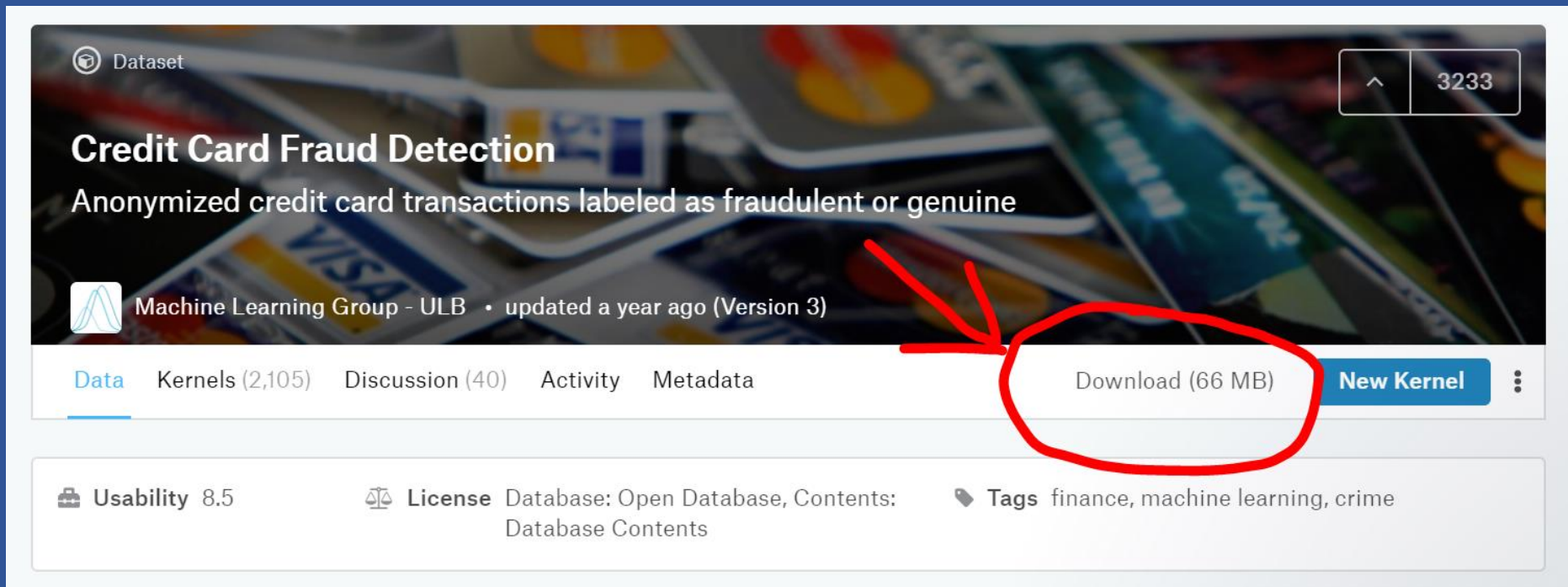
<https://raw.githubusercontent.com/laploy/ML.NET/master/Credit-Card/creditCard-azure-model.jpg>



Question and Data

Question: Is this credit card transaction a fraud?

Dataset: <https://www.kaggle.com/mlg-ulb/creditcardfraud>



The screenshot shows the Kaggle dataset page for 'Credit Card Fraud Detection'. The page features a header with the dataset title and a description: 'Anonymized credit card transactions labeled as fraudulent or genuine'. Below the header, there is a navigation bar with tabs for 'Data', 'Kernels (2,105)', 'Discussion (40)', 'Activity', and 'Metadata'. A red circle highlights the 'Download (66 MB)' button, with a red arrow pointing to it from the left. To the right of the download button is a 'New Kernel' button. Below the navigation bar, there is a section with 'Usability 8.5', 'License Database: Open Database, Contents: Database Contents', and 'Tags finance, machine learning, crime'.

Dataset

Credit Card Fraud Detection
Anonymized credit card transactions labeled as fraudulent or genuine

Machine Learning Group - ULB • updated a year ago (Version 3)

Data Kernels (2,105) Discussion (40) Activity Metadata

Download (66 MB) New Kernel

Usability 8.5 License Database: Open Database, Contents: Database Contents Tags finance, machine learning, crime

Dataset description

- Transactions made by credit cards in September 2013
- European cardholders
- Occurred in two days
- 284,807 rows, 31 columns
- Normal Class = 0, fraud = 1
- 492 frauds

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | AA | AB | AC | AD | AE |
|----|------|----|----|-----|-----|-----|-----|-----|------|------|------|-----|-----|------|------|-----|-----|------|------|------|-----|------|------|------|-----|-----|-----|------|------|--------|-------|
| 1 | Time | V1 | V2 | V3 | V4 | V5 | V6 | V7 | V8 | V9 | V10 | V11 | V12 | V13 | V14 | V15 | V16 | V17 | V18 | V19 | V20 | V21 | V22 | V23 | V24 | V25 | V26 | V27 | V28 | Amount | Class |
| 2 | 0 | -1 | -0 | 2.5 | 1.4 | -0 | 0.5 | 0.2 | 0.1 | 0.36 | 0.09 | -1 | -1 | -1 | -0.3 | 1.5 | -0 | 0.2 | 0.03 | 0.4 | 0.3 | -0 | 0.28 | -0.1 | 0.1 | 0.1 | -0 | 0.13 | -0 | 149.62 | 0 |
| 3 | 0 | 1 | 0 | 0.2 | 0.4 | 0.1 | -0 | -0 | 0.1 | -0.3 | -0.2 | 2 | 1.1 | 0.49 | -0.1 | 0.6 | 0.5 | -0.1 | -0.2 | -0.1 | -0 | -0.2 | -0.6 | 0.1 | -0 | 0.2 | 0.1 | -0 | 0 | 2.69 | 0 |
| 4 | 1 | -1 | -1 | 1.8 | 0.4 | -1 | 1.8 | 0.8 | 0.2 | -1.5 | 0.21 | 1 | 0.1 | 0.72 | -0.2 | 2.3 | -3 | 1.1 | -0.1 | -2.3 | 0.5 | 0.25 | 0.77 | 0.9 | -1 | -0 | -0 | -0.1 | -0.1 | 378.66 | 0 |
| 5 | 1 | -1 | -0 | 1.8 | -1 | -0 | 1.2 | 0.2 | 0.4 | -1.4 | -0.1 | -0 | 0.2 | 0.51 | -0.3 | -1 | -1 | -0.7 | 1.97 | -1.2 | -0 | -0.1 | 0.01 | -0.2 | -1 | 0.6 | -0 | 0.06 | 0.1 | 123.5 | 0 |
| 6 | 2 | -1 | 1 | 1.5 | 0.4 | -0 | 0.1 | 0.6 | -0.3 | 0.82 | 0.75 | -1 | 0.5 | 1.35 | -1.1 | 0.2 | -0 | -0.2 | -0 | 0.8 | 0.4 | -0 | 0.8 | -0.1 | 0.1 | -0 | 0.5 | 0.22 | 0.2 | 69.99 | 0 |
| 7 | 2 | -0 | 1 | 1.1 | -0 | 0.4 | -0 | 0.5 | 0.3 | -0.6 | -0.4 | 1 | 0.4 | -0.4 | -0.1 | 0.5 | 0.4 | -0.1 | 0.07 | -0 | 0.1 | -0.2 | -0.6 | -0 | -0 | -0 | 0.1 | 0.25 | 0.1 | 3.67 | 0 |
| 8 | 4 | 1 | 0 | 0 | 1.2 | 0.2 | 0.3 | -0 | 0.1 | 0.46 | -0.1 | -1 | -0 | -0.8 | 0.17 | 0.1 | -0 | 0 | -0.6 | -0 | -0 | -0.2 | -0.3 | -0.2 | -1 | 0.8 | -0 | 0.03 | 0 | 4.99 | 0 |
| 9 | 7 | -1 | 1 | 1.1 | -0 | 0.9 | 0.4 | 1.1 | -3.8 | 0.62 | 1.25 | -1 | 0.3 | 1.76 | -1.3 | 0.7 | -0 | -1.2 | -0.4 | 0.32 | -0 | 1.94 | -1 | 0.1 | -1 | -0 | -0 | -1.2 | -1.1 | 40.8 | 0 |
| 10 | 7 | -1 | 0 | -0 | -0 | 2.7 | 3.7 | 0.4 | 0.9 | -0.4 | -0.4 | -1 | -0 | -0.3 | 0.07 | -0 | -0 | -0.5 | 0.12 | 0.57 | 0.1 | -0.1 | -0.3 | -0.2 | 1 | 0.4 | -0 | 0.01 | 0.1 | 93.2 | 0 |
| 11 | 9 | -0 | 1 | 1 | -0 | 0.5 | -0 | 0.7 | 0.1 | -0.7 | -0.4 | 1 | 0.8 | 1.01 | -0.4 | 0.2 | 0.7 | -0.5 | 0.48 | 0.45 | 0.2 | -0.2 | -0.6 | -0.1 | -0 | -0 | 0.1 | 0.25 | 0.1 | 3.68 | 0 |

Create Data Set

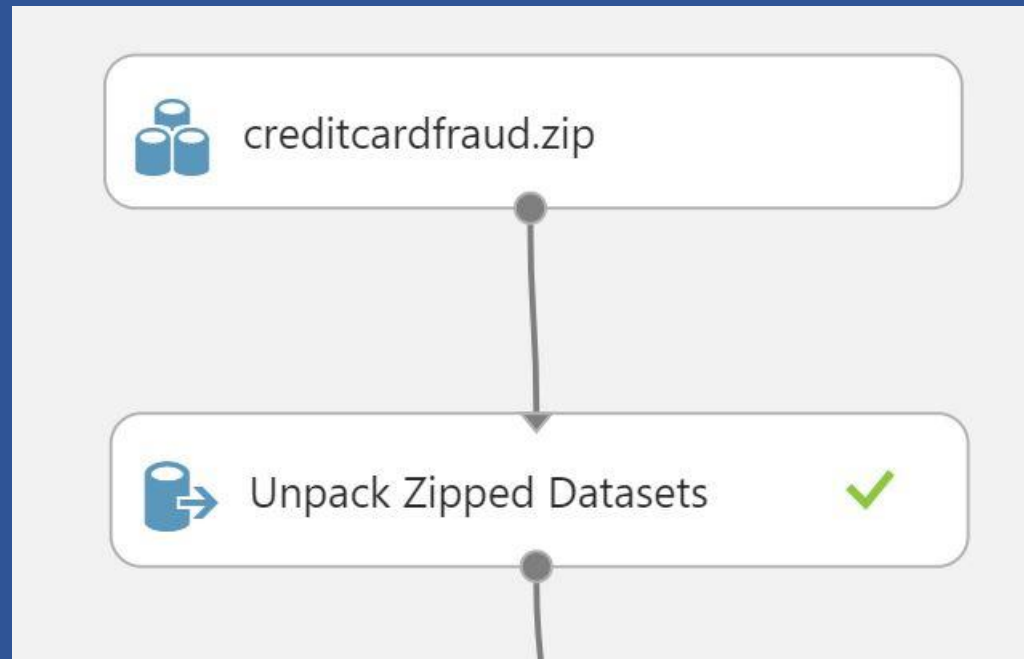
Click DATASET / + NEW / import both datasets

datasets

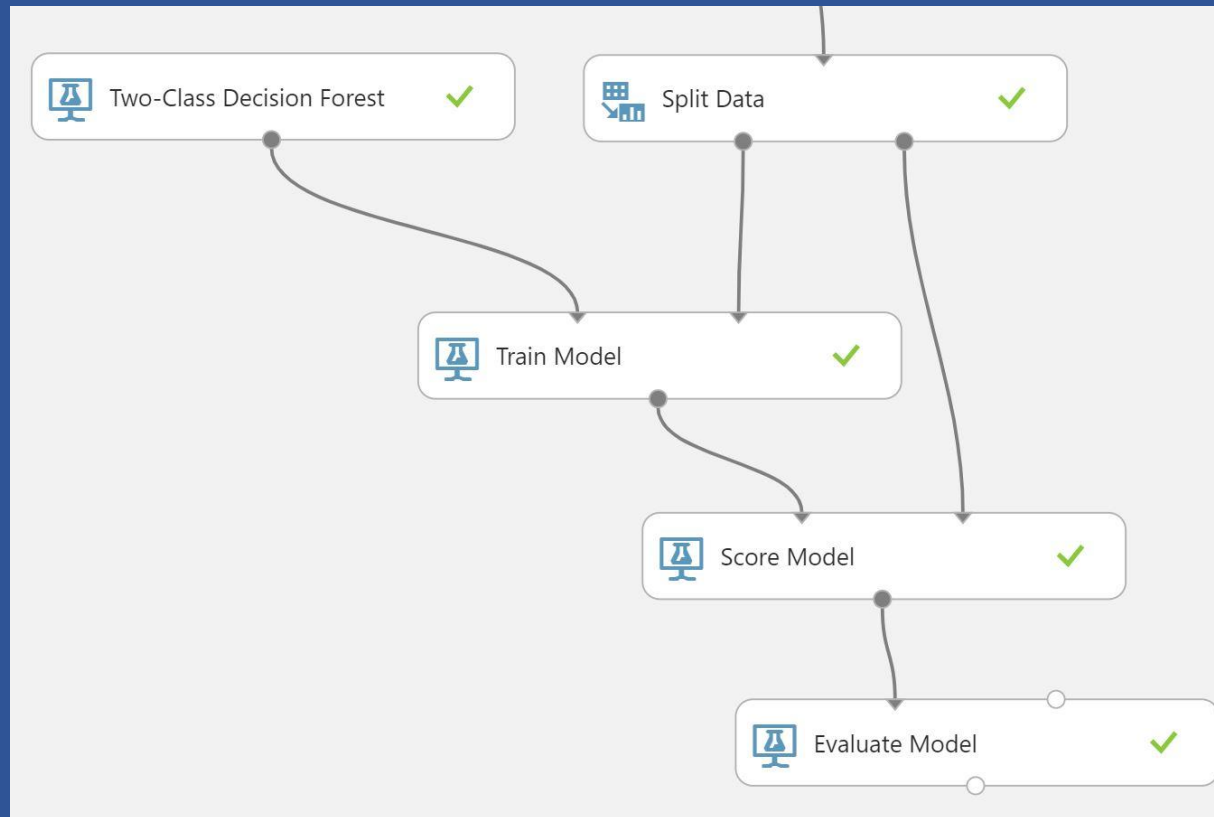
MY DATASETS SAMPLES

| | NAME | SUBMIT... |
|--------------------------|---------------------|-----------|
| <input type="checkbox"/> | creditcardfraud.zip | laploy |

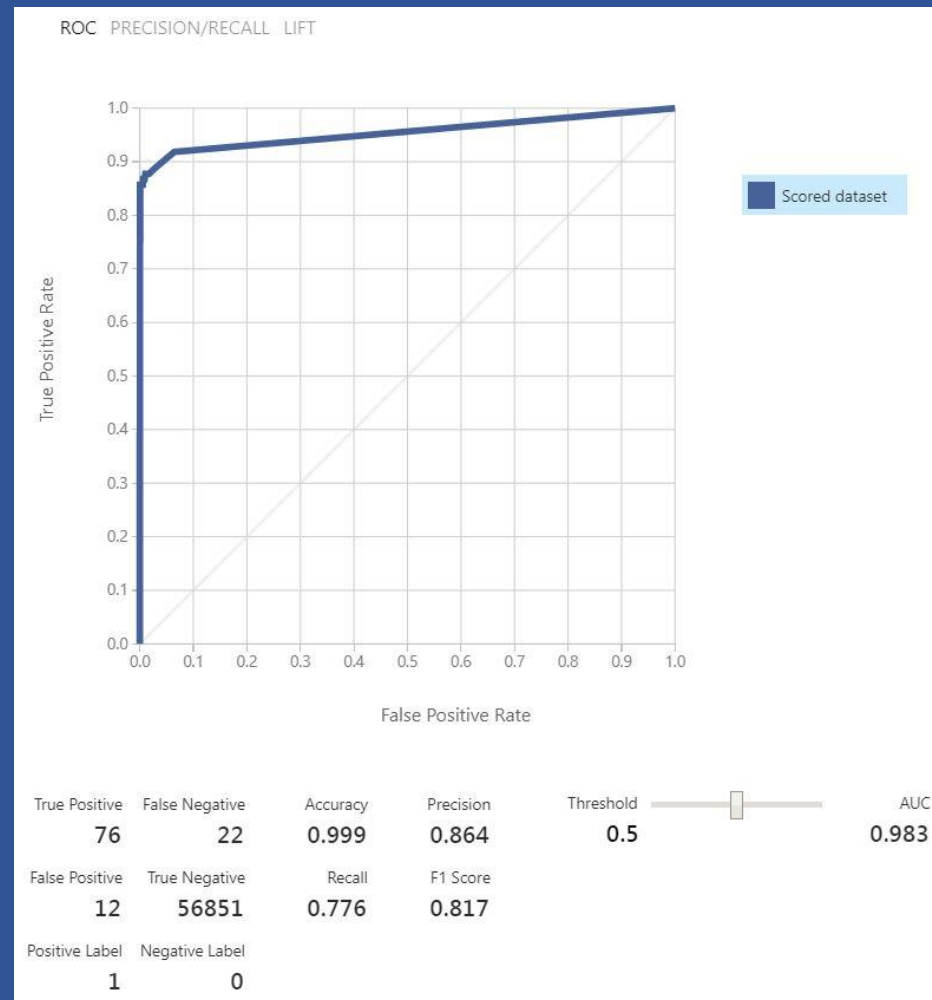
Add Datasets, Unzip



Split, train, score, and evaluate



Score result



What next?

Credit card AutoML