

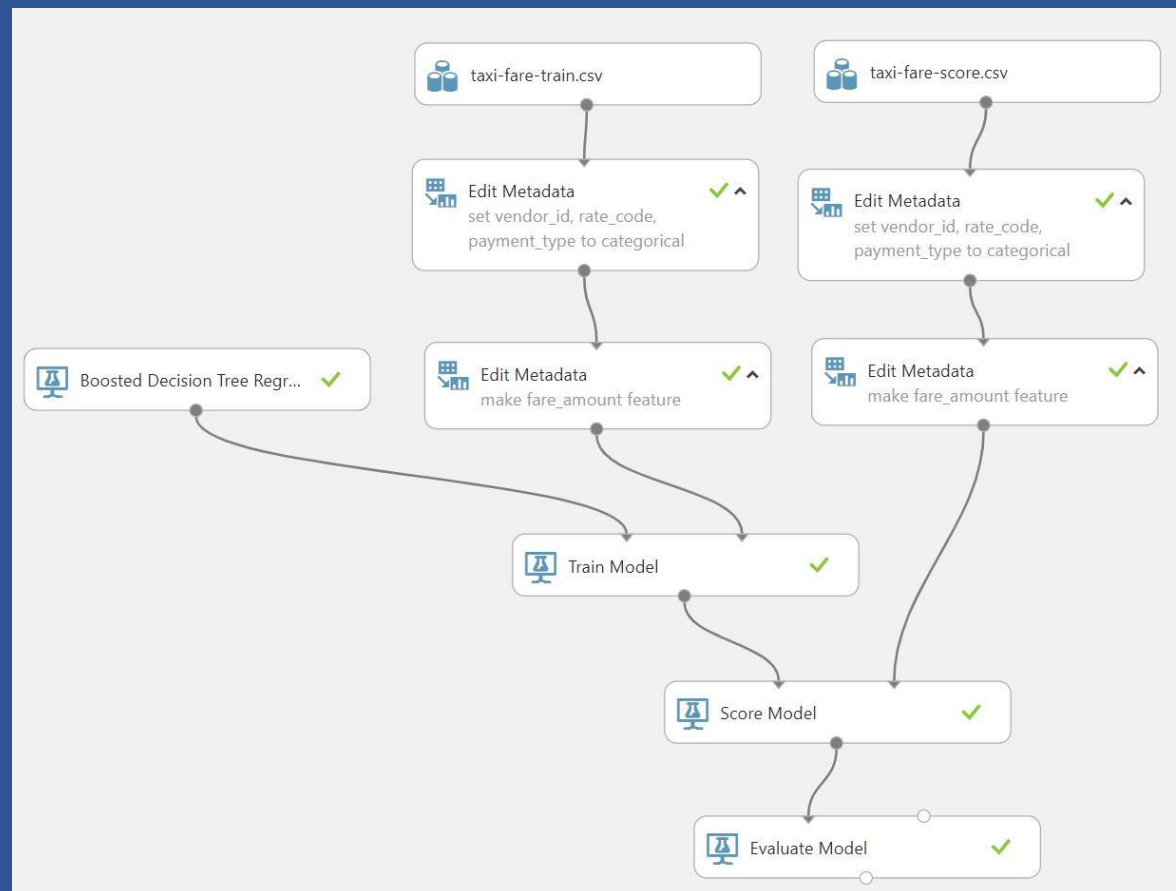
Taxi fair in Azure ML

In this session

- Question and Data
- Dataset description
- Create Data Set
- Place dataset
- Drop column ID
- Specify Label
- Hash feature
- Drop area and description
- Split Data
- Train, Score, Evaluate
- Metrics for multiclass classification Evaluation

The finished model

<https://raw.githubusercontent.com/laploy/ML.NET/master/Taxi-fare/taxi-fare-model.JPG>



Question and Data

Question: How much is the taxi fair?

Dataset:

taxi-fare-train.csv

<https://raw.githubusercontent.com/laploy/ML.NET/master/Taxi-fare/taxi-fare-train.csv>

taxi-fare-score.csv

<https://raw.githubusercontent.com/laploy/ML.NET/master/Taxi-fare/taxi-fare-score.csv>

taxi-fare-batch.csv

<https://raw.githubusercontent.com/laploy/ML.NET/master/Taxi-fare/taxi-fare-batch.csv>

Dataset description

vendor_id: A code indicating the TPEP provider that provided the record.

rate_code: The final rate code in effect at the end of the trip.

1. Standard rate
2. JFK
3. Newark
4. Nassau or Westchester
5. Negotiated fare
6. Group ride

passenger_count: The number of passengers in the vehicle

trip_time_in_secs:

trip_distance: The elapsed trip distance in miles reported by the taximeter.

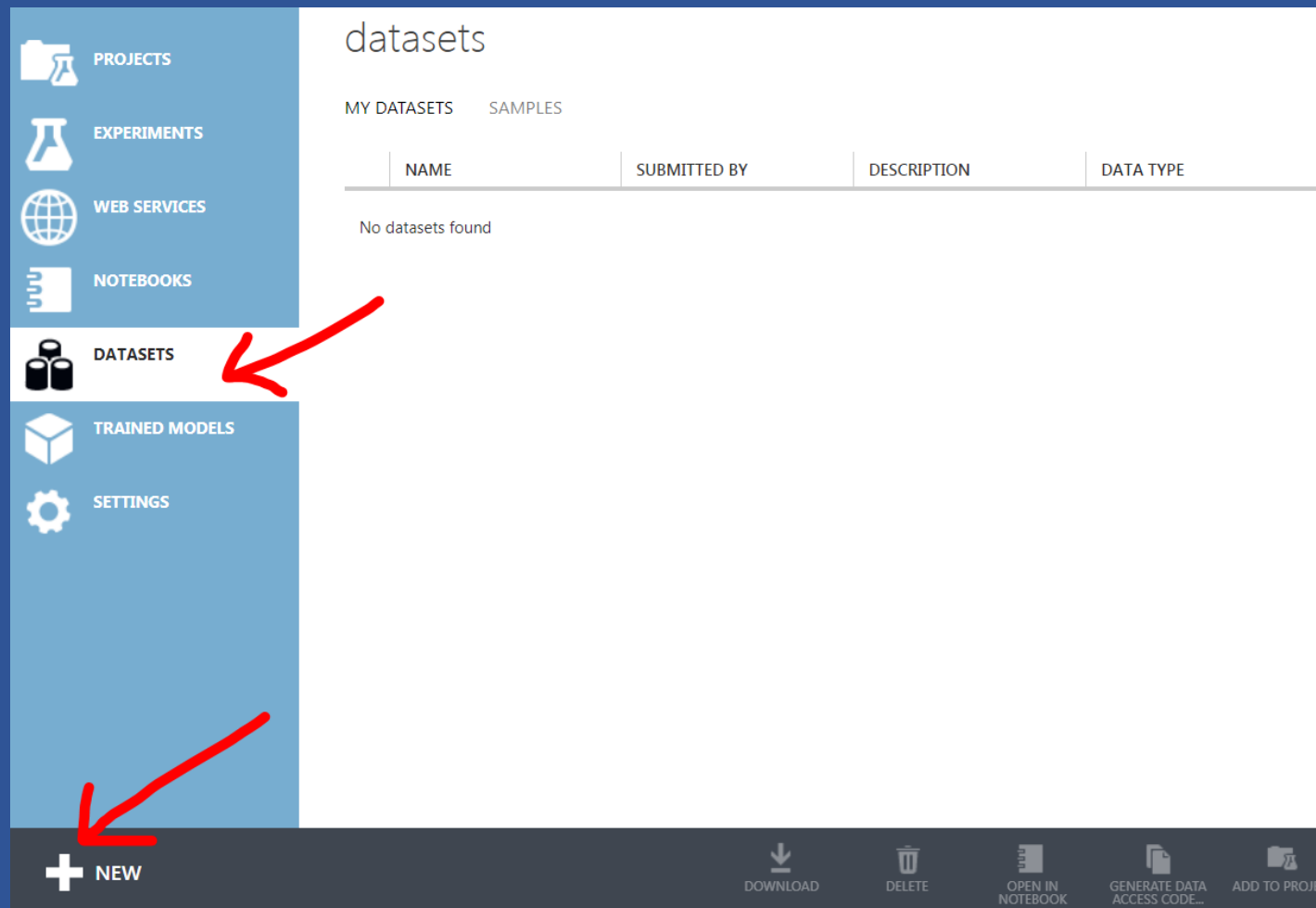
payment_type: A numeric code signifying how the passenger paid for the trip.

1= Credit card 2= Cash 3= No charge 4= Dispute 5= Unknown 6= Voided trip

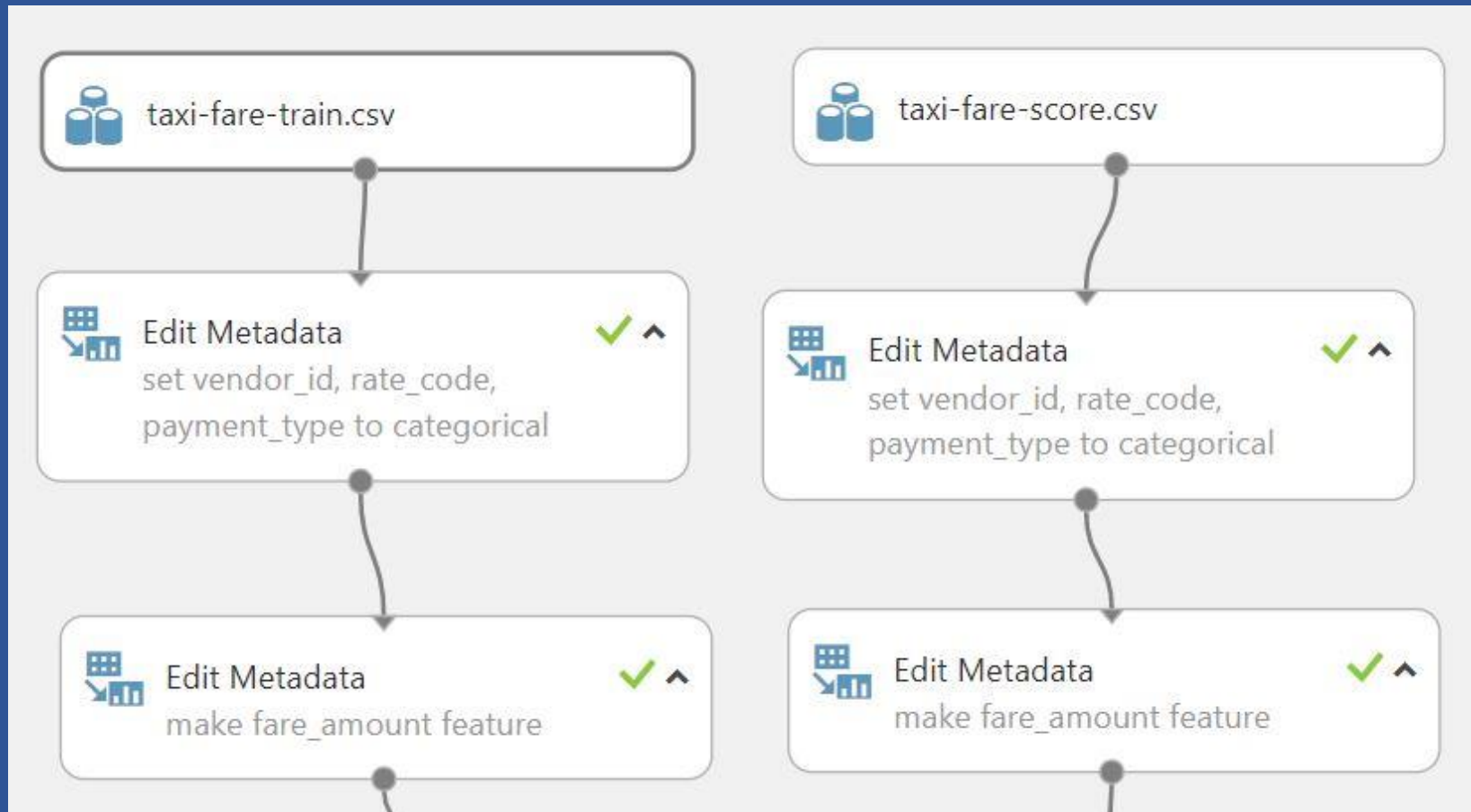
fare_amount: The time-and-distance fare calculated by the meter.

Create Data Set

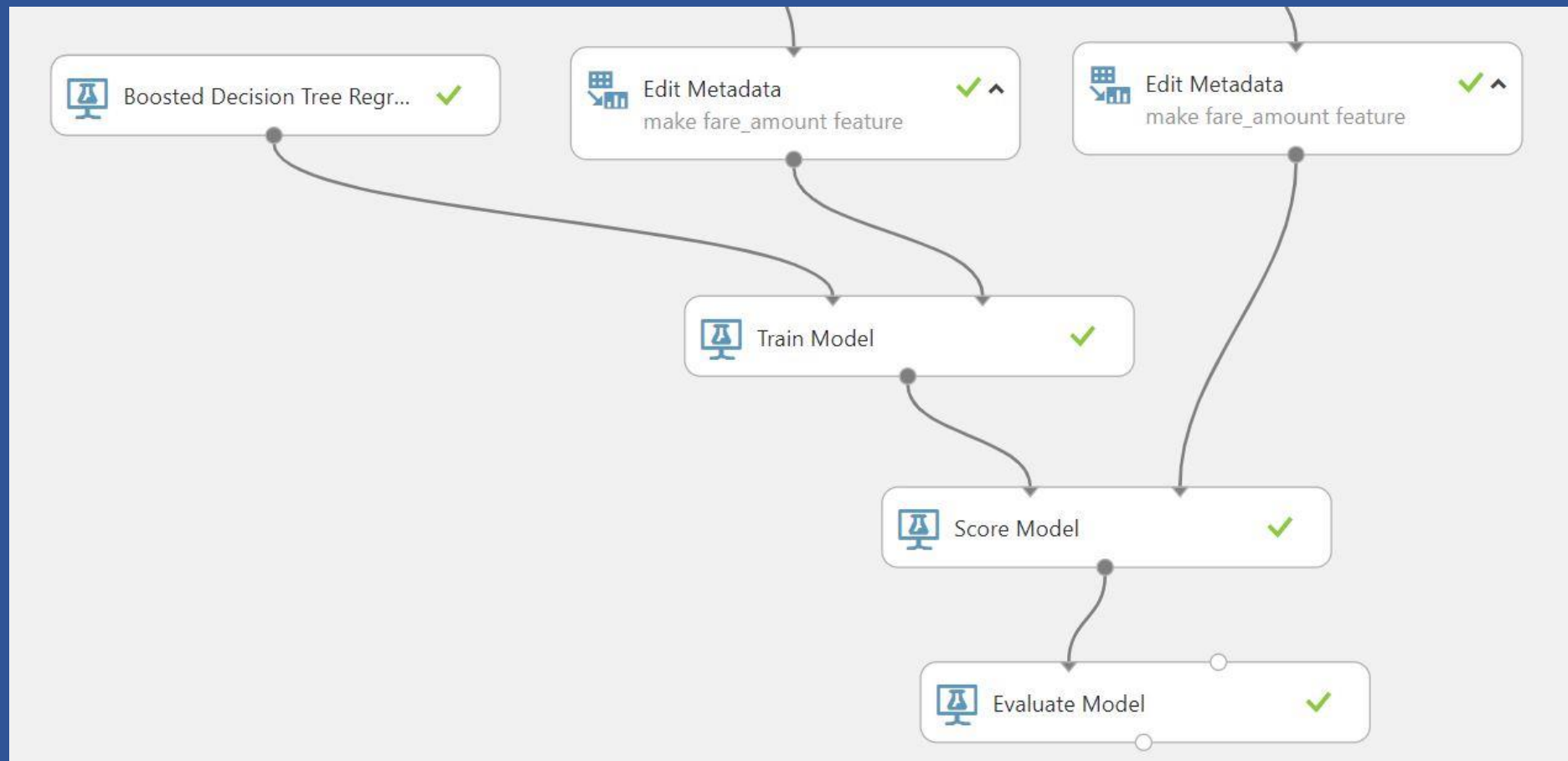
Click DATASET / + NEW / import both datasets



Add Datasets, set categorical and make feature



Train, Score and Evaluate



Evaluation Metrics

Mean absolute error (MAE)

average of all the model errors, where model error is the distance between the predicted label value and the correct label value.

Coefficient of determination

how well data fits a model. Ranges from 0 to 1. A value of 0 means that the data is random or otherwise cannot be fit to the model. A value of 1 means that the model exactly matches the data.

What next?

Taxi fare AutoML