Taxi Fare

Write Code

(Regression-Fast Tree)

What's in this session?

1. Question and Data
2. Create project
3. Add NuGet packages
4. Add using name space
5. Create data set input/output scheme
6. Set data set path
7. Load data

8. Add algorithm
9. Train the model
10. Save model
11. Evaluate the model and show accuracy stats
12. Predict single item

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

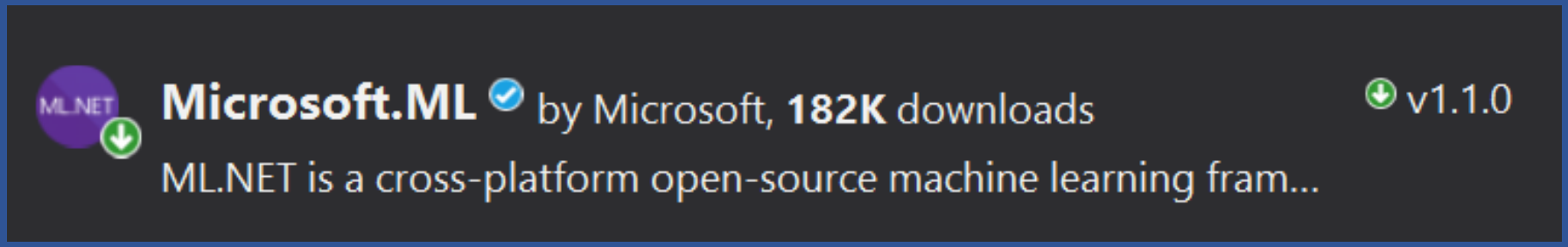
fare_amount: The time-and-distance fare calculated by the meter.

Create New Project

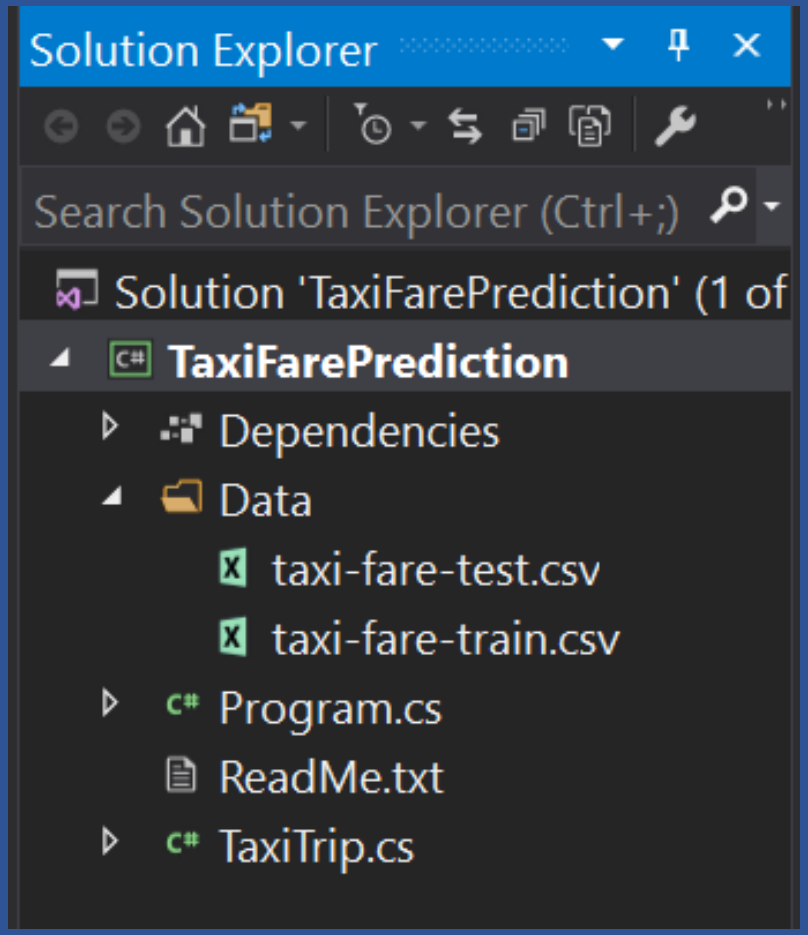
Create new .NET CORE console app project name = “Taxi”

Add NuGet Package

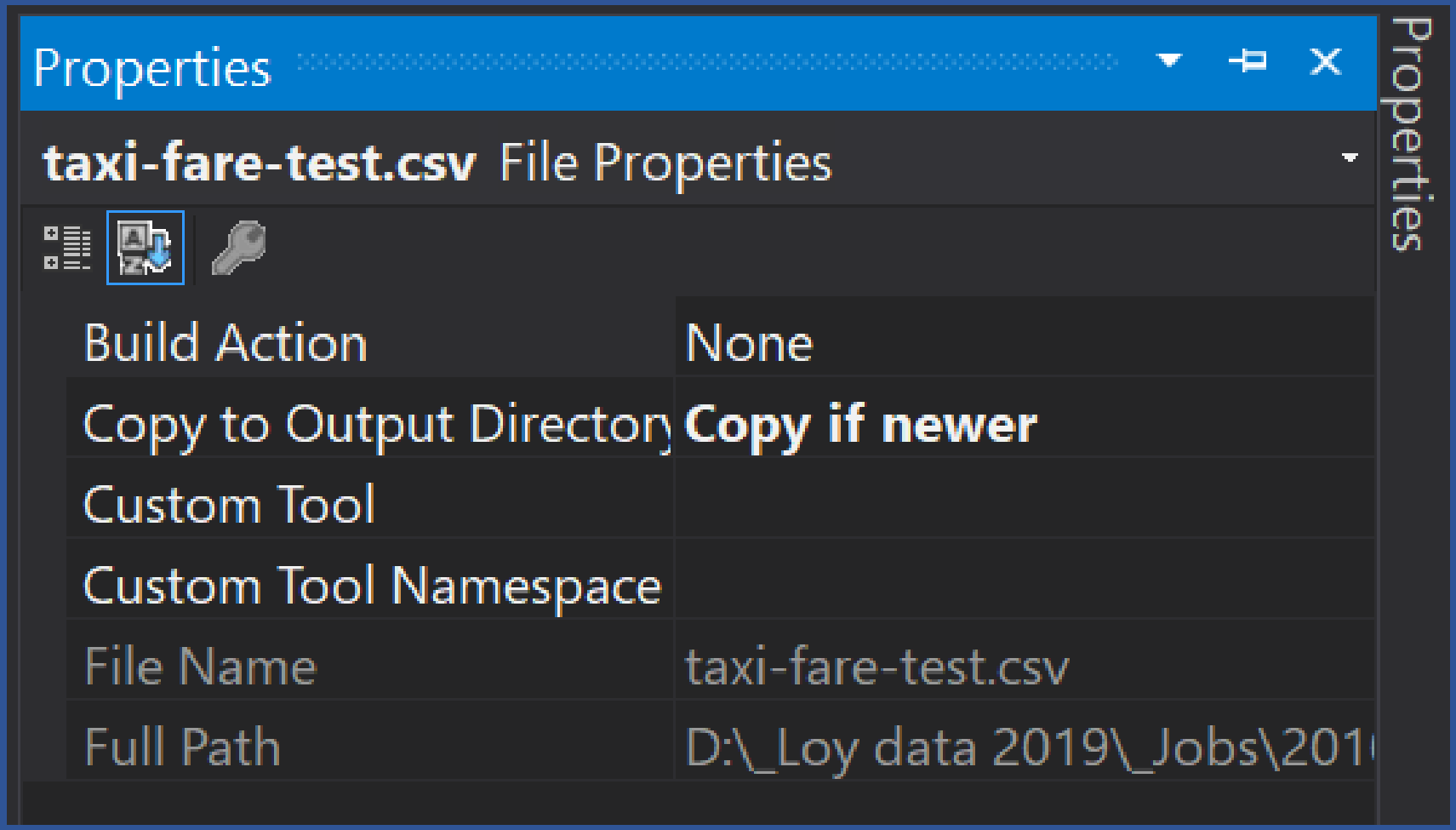
- Microsoft.ML
- Microsoft.ML.Fast



Add new folder “Data”
Copy datasets to this folder



Set property of each datasets to “Copy if newer”



Create data set input/output scheme

```
 8   namespace TaxiFarePrediction
 9   {
10       // input data class
11       public class TaxiTrip...
28
29       // output data class
30       public class TaxiTripFarePrediction
31       {
32           // In case of the regression task the Score column contai
33           [ColumnName("Score")]
34           // Use the float type to represent floating-point values
35           public float FareAmount;    // is the prediction output
36       }
37   }
```

Set data set paths

```
class Program
{
    static readonly string _trainDataPath = Path.Combine
        (Environment.CurrentDirectory, "Data", "taxi-fare-train.csv");
    static readonly string _testDataPath = Path.Combine
        (Environment.CurrentDirectory, "Data", "taxi-fare-test.csv");
```

Create ML Context, Load data, and Process Data

```
static void Main(string[] args)
{
    // create comtext
    MLContext mlContext = new MLContext(seed: 0);

    // train model
    var model = Train(mlContext, _trainDataPath);

    // Evaluate the model
    Evaluate(mlContext, model);

    // Use the model for predictions
    TestSinglePrediction(mlContext, model);


}
```

The program output result

```
Load train data.
create pipeline.
Strat training. 6/16/2019 9:35:31 AM
Training done. 6/16/2019 9:35:39 AM
Loads the test dataset.
Creates the regression evaluator.
Evaluates the model and creates metrics.

*************************************************
* Model quality metrics evaluation
*------------------------------------------------
* RSquared Score: 0.92
* Root Mean Squared Error: 2.81
*************************************************
Predicted fare: 15.7855, actual fare: 15.5
*************************************************
```

What's next?

Predict taxi fare