Movie

Write Code ML.NET

(Matrix factorization)

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 movie interesting to the user?

Dataset:

Train

https://raw.githubusercontent.com/laploy/ML.NET/master/Movie/recommendation-ratings-train.csv

Test

https://raw.githubusercontent.com/laploy/ML.NET/master/Movie/recommendation-ratings-test.csv

Title

https://raw.githubusercontent.com/laploy/ML.NET/master/Movie/recommendation-movies.csv

Dataset description

- User Id Feature
- moiveId Feature
- rating Label
- timestamp non-Feature

26,772 users
227,472 rows

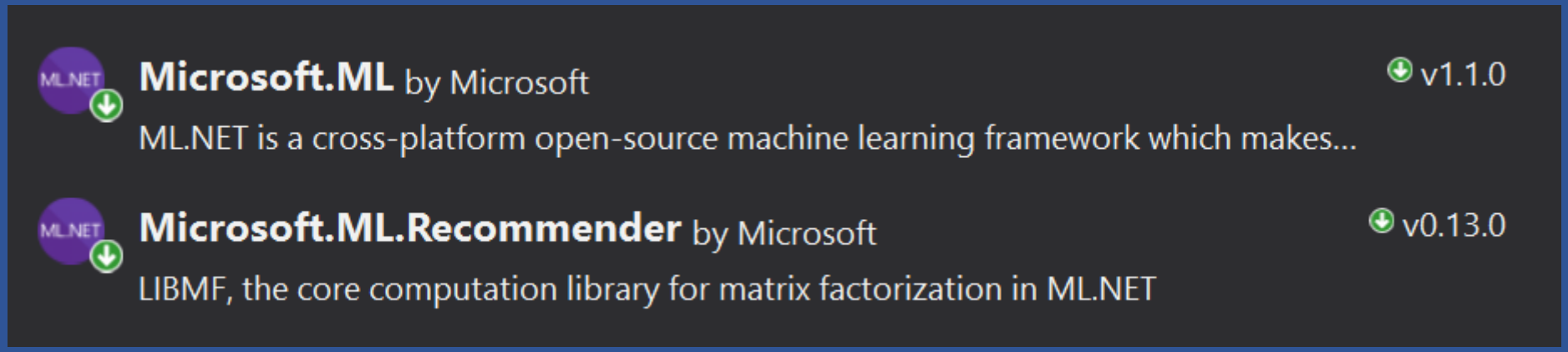
	A	B	C	D
1	UserId	MovieId	Rating	Timestamp
2	1	68646	10	1381620027
3	1	113277	10	1379466669
4	2	454876	8	1394818630
5	2	790636	7	1389963947
6	2	816711	8	1379963769
7	2	1091191	7	1391173869
8	2	1322269	7	1391529691
9	2	1433811	8	1380453043

Create New Project

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

Add NuGet Package

- Microsoft.ML
- Microsoft.ML.Recommender



Create data scheme and helper class

```
6 references
public class MovieRating...

1 reference
class MovieRatingPrediction...

9 references
class Movie...
```

Write code in class Program

```
1   using System;
2   using Microsoft.ML;
3   using Microsoft.ML.Trainers;
4
5   namespace Movie
6   {
        0 references
7       class Program
8       {
9           private static string TrainingDataLocation =
10              @"E:\ml\recommendation-ratings-train.csv";
11          private static string TestDataLocation =
12              @"E:\ml\recommendation-ratings-test.csv";
13
14          private const float predictionuserId = 6;
15          private const int predictionmovieId = 10;
16
            0 references
17          static void Main(string[] args)...
98      }
99  }
100
```

The program output result

```
=============== Training the model ===============
iter      tr_rmse          obj
   0       1.5055   2.8315e+05
   1       0.9193   1.4063e+05
   2       0.8677   1.3077e+05
   3       0.8427   1.2618e+05
   4       0.8278   1.2361e+05
   5       0.8143   1.2141e+05
   6       0.8007   1.1945e+05
   7       0.7873   1.1777e+05
   8       0.7728   1.1602e+05
   9       0.7571   1.1428e+05
  10       0.7421   1.1261e+05
  11       0.7279   1.1130e+05
  12       0.7157   1.1009e+05
  13       0.7029   1.0891e+05
  14       0.6916   1.0779e+05
  15       0.6810   1.0690e+05
  16       0.6707   1.0607e+05
  17       0.6616   1.0531e+05
  18       0.6526   1.0452e+05
  19       0.6442   1.0396e+05
=============== Evaluating the model ===============
The model evaluation metrics RootMeanSquaredError:0.9643936856269949

=============== Make a single movie rating prediction ===============
For userId:6 movie rating prediction (1 - 5 stars) for movie:GoldenEye (1995) is:3.7
```

What's next?