

Transformer

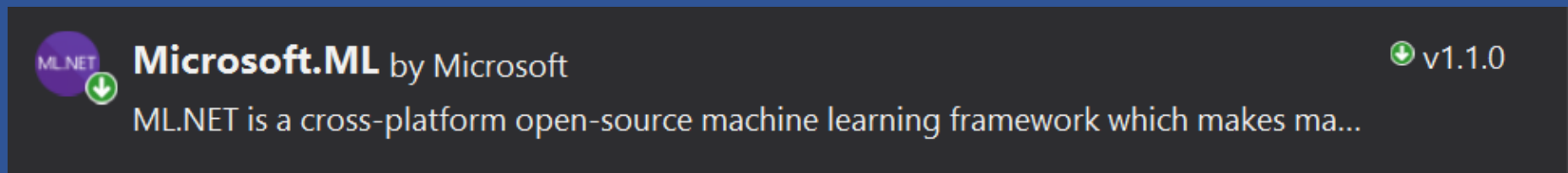
What's in this session?

- Split data
- Features Vector
- Features Vector
- Drop column
- Default column

Create New Project

Create new .NET CORE console app project name = "Preview"

Add NuGet Packages



Create data model via Classes

```
public partial class Adult
{
    [LoadColumn(0)]
    public short Id { get; set; }

    [LoadColumn(1)]
    public byte Age { get; set; }

    [LoadColumn(2)]
    public string Workclass { get; set; }

    [LoadColumn(3)]
    public string Education { get; set; }
}
```

Split data

```
DataOperationsCatalog.TrainTestData dataSplit = mlContext.Data.TrainTestSplit(  
    data,  
    testFraction: 0.2);  
IDataView trainData = dataSplit.TrainSet;  
IDataView testData = dataSplit.TestSet;
```

```
2      50      Sales  
7      49      Private  
8      52      Programmer  
16     34      Private  
17     25      Self-emp-not-inc  
18     32      Teacheer  
19     38      Private  
23     35      Student  
26     56      Local-gov  
27     19      Private  
31     23      Local-gov  
34     30      Federal-gov  
35     22      State-gov  
  
-----show all row-----  
1, 2, 7, 8, 16, 17, 18, 19, 23, 26, 27, 31, 34, 35,  
  
-----Train data-----  
1, 2, 7, 8, 16, 19, 23, 26, 27, 31, 34, 35,  
  
-----Test data-----  
17, 18,
```

Features Vector

Create data class

```
public class IrisData
{
    [LoadColumn(0)]
    public float SepalLength;

    [LoadColumn(1)]
    public float SepalWidth;

    [LoadColumn(2)]
    public float PetalLength;

    [LoadColumn(3)]
    public float PetalWidth;
}
```

Features Vector

```

IEstimator<ITransformer> dataPrepEstimator =
mlContext.Transforms.Concatenate("Features", "SepalLength", "SepalWidth")
.Append(mlContext.Transforms.NormalizeMinMax("Features"));
// Create data prep transformer
ITransformer dataPrepTransformer = dataPrepEstimator.Fit(data);
// Apply tranforms to training data
IDataView transformedTrainingData = dataPrepTransformer.Transform(data);

```

```

-----show original data-----
SepalLength: Single | SepalWidth: Single | PetalLength: Single | PetalWidth: Single |
-----
5.3      3.7      1.5
5        3.3      1.4
5.1      2.5      3
5.7      2.8      4.1
6.2      3.4      5.4
5.9      3        5.1

-----show Concatenate-----
5.3, 3.7,
5, 3.3,
5.1, 2.5,
5.7, 2.8,
6.2, 3.4,
5.9, 3,

-----show Concatenate & Normalize-----
0.8548388, 0.99999994,
0.8064517, 0.89189184,
0.8225807, 0.67567563,
0.91935486, 0.7567567,
1, 0.9189189,
0.95161295, 0.8108108,

```

Drop column

Create data class

```
[LoadColumn(0)]  
public float Size { get; set; }  
[LoadColumn(1)]  
public float Price { get; set; }
```


Drop column

```
// drop a column  
var d = mlContext.Transforms.DropColumns("Size");
```

```
-----show original data-----  
Size: Single | Price: Single |  
-----  
1.01      7366  
0.49      985  
0.31      544  
1.51     9140  
0.37      493  
0.73     3011  
1.53    11413  
  
-----show original data-----  
Price: Single |  
-----  
7366  
985  
544  
9140  
493  
3011  
11413
```

Default column

```
// change feature and lable from default column names
var model = mlContext.Regression.Trainers.Sdca(
    ...
    labelColumnName: "Cost",
    featureColumnName: "Karat"
);
```

```
-----show original data-----
Size: Single | Price: Single |
-----
```

1.01	7366
0.49	985
0.31	544
1.51	9140
0.37	493
0.73	3011
1.53	11413

```
-----show original data-----
Karat: Single | Cost: Single |
-----
```

1.01	7366
0.49	985
0.31	544
1.51	9140
0.37	493
0.73	3011
1.53	11413

What's next?

