

Spam

Write Code ML.NET

(Multiclass Classification)

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: This message is normal or a spam?

Dataset:

Train

<https://raw.githubusercontent.com/laploy/ML.NET/master/Spam/spam-train.tsv>

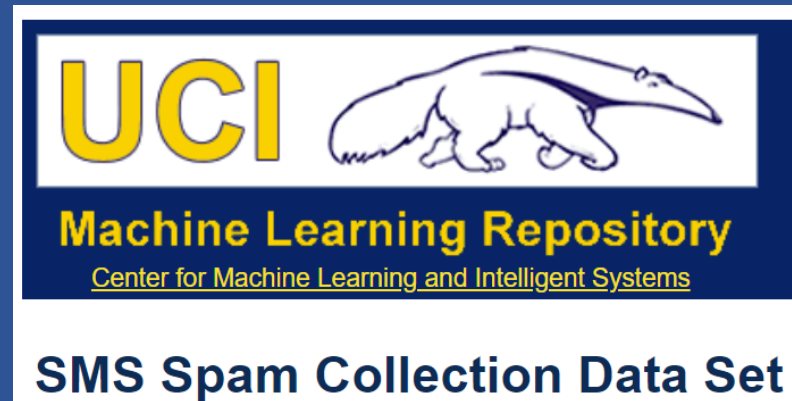
Score

<https://raw.githubusercontent.com/laploy/ML.NET/master/Spam/spam-score.tsv>

Test

<https://raw.githubusercontent.com/laploy/ML.NET/master/Spam/spam-test.tsv>

Dataset description



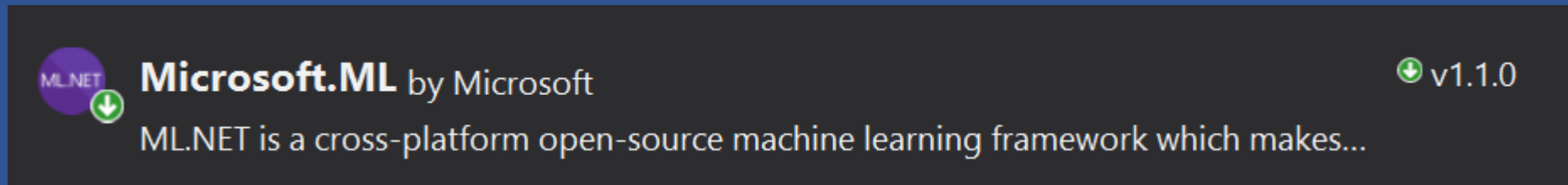
class	message
spam	REMINDER FROM O2: To get 2.50 pounds free cal
ham	Rofl. Its true to its name
spam	This is the 2nd time we have tried 2 contact u. U h
ham	Will ã¼ b going to esplanade fr home?
ham	Pity, * was in mood for that. So...any other sugges
spam	Want explicit SEX in 30 secs? Ring 02073162414 n
ham	The guy did some bitching but I acted like i'd be i

Create New Project

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

Add NuGet Package

- Microsoft.ML



Create data scheme and helper class

```
4  namespace Spam
5  {
6      class SpamInput...
13
14     class SpamPrediction...
19  }
20
```

Download helper class from github

<https://github.com/laploy/ML.NET/blob/master/Spam/ConsoleHelper.cs>

Write code in class Program

```
class Program
{
    private static string trainDataPath = @"E:\ML\spam-train.tsv";
    private static string saveModelPath = @"E:\ML\spam-model.zip";

    static void Main(string[] args)...

    public static void ClassifyMessage(PredictionEngine<SpamInput, S
    {
        var input = new SpamInput { Message = message };
        var prediction = predictor.Predict(input);

        Console.WriteLine("The message '{0}' is {1}",
            input.Message,
            prediction.isSpam == "spam" ? "spam" : "not spam");
    }
}
```


The program output result

```
===== Cross-validating to get model's accuracy metrics =====
*****
*      Metrics for Microsoft.ML.Data.EstimatorChain`1[Microsoft.ML.Transforms.KeyToValueMappingTransformer]
*      Classification model
*      -----
*      Average MicroAccuracy:    0.984 - Standard deviation: (.003) - Confidence Interval 95%: (.003)
*      Average MacroAccuracy:    0.961 - Standard deviation: (.012) - Confidence Interval 95%: (.011)
*      Average LogLoss:          .059 - Standard deviation: (.009) - Confidence Interval 95%: (.009)
*      Average LogLossReduction: .85  - Standard deviation: (.027) - Confidence Interval 95%: (.026)
*****
===== Predictions for below data=====
The message 'That's a great idea. It should work.' is not spam
The message 'free medicine winner! congratulations' is spam
The message 'Yes we should meet over the weekend!' is not spam
The message 'you win pills and free entry vouchers' is spam
===== End of process, hit any key to finish =====
```

What's next?

