

Sentiment AutoML

What's in this session?

1. Install ML.NET Model Builder
2. Create new .NET CORE console project and add Machine Learning job
3. Pick a Scenario / Price Prediction
4. Set Data File
5. Train 60 seconds
6. Understand Train result
7. Understand evaluation result
8. Generate Code
9. Examine Code

Install ML.NET Model Builder

<https://marketplace.visualstudio.com/items?itemName=MLNET.07>



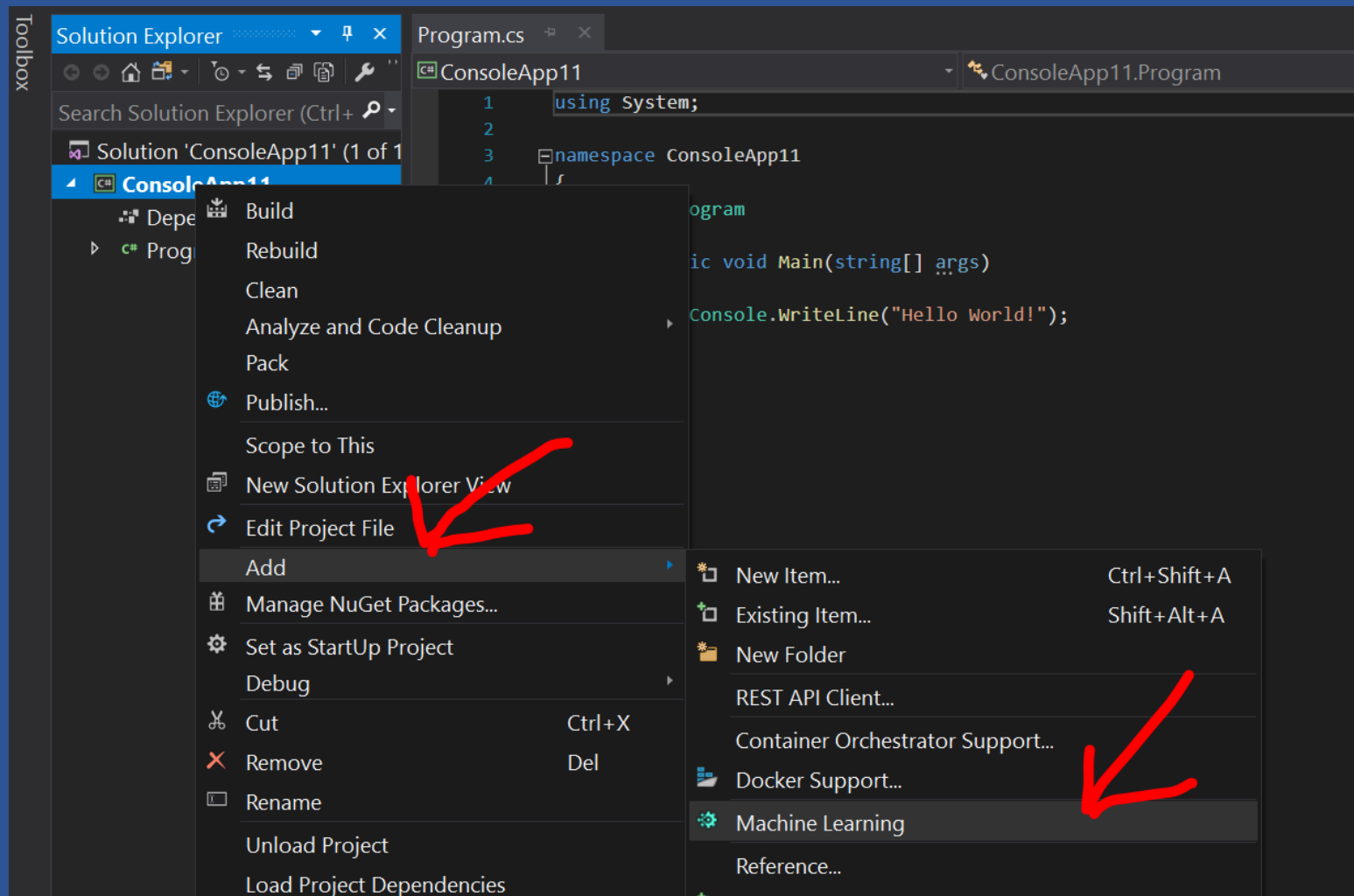
ML.NET Model Builder (Preview)

Microsoft | 24,428 installs |  39,194 downloads | ★★★★★ (12) | Free

Simple UI tool to build custom machine learning models.

Download

Create new .NET CORE console project and add Machine Learning



Pick a Scenario / Sentiment Analysis



Classify data into 2 categories (binary classification), e.g. predict positive or negative sentiment of comments.

Set Data File

Data / File / yelp_labelled1.tsv
Label column name = Label

Select a file:

E:\ml\yelp_labelled1.tsv ...

Supported file formats: csv or tsv. Maximum file size: 1 GB.

Column to Predict (Label): ⓘ

Label

Data Preview


Message	Label (Label)
Wow... Loved this place.	1
Crust is not good.	0
Not tasty and the texture was just nasty.	0
Stopped by during the late May bank holiday off Rick Steve recommendation and loved it.	1
The selection on the menu was great and so were the prices.	1
Now I am getting angry and I want my damn pho.	0
Honeslty it didn't taste THAT fresh.)	0

Train 60 seconds

Train

Specify a time to train for evaluating various models.
How long should I train for?

Input

Time to train (seconds): 

Start training

Progress

Start training to see progress and results

Status:

Best Accuracy:

Best Algorithm:

Last Algorithm:

Understand Train result

Progress

Start training to see progress and results

Status:	Done
Best Accuracy:	87.36%
Best Algorithm:	SdcaLogisticRegressionBinary
Last Algorithm:	SgdCalibratedBinary

Understand evaluation result

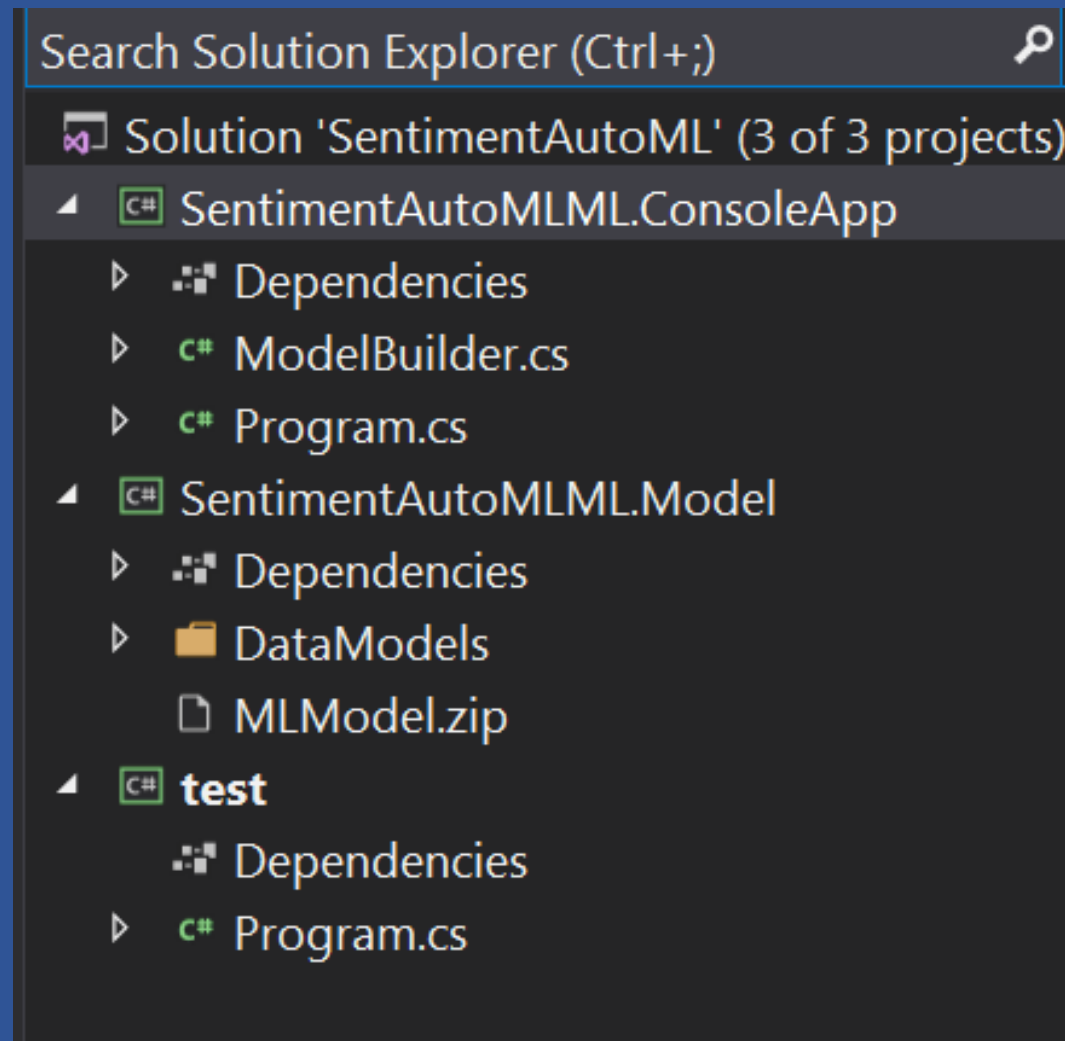
Output

ML Task: binary-classification
Dataset: yelp_labelled1.tsv
Column to Predict (Label): Label
Best Model: SdcaLogisticRegressionBinary
Best Model Accuracy: 87.36%
Training Time: 60.35 seconds
Models Explored (Total): 150

Top 5 models explored

Rank	Trainer	Accuracy	AUC	AUPRC	F1-score	Duration
1	SdcaLogisticRegressionBinary	0.8736	0.9355	0.9319	0.8791	0.1
2	SdcaLogisticRegressionBinary	0.8736	0.9392	0.9373	0.8866	0.1
3	SgdCalibratedBinary	0.8736	0.9355	0.9393	0.8889	0.1
4	SdcaLogisticRegressionBinary	0.8736	0.9308	0.9242	0.8791	0.8
5	LbfgsLogisticRegressionBinary	0.8736	0.9323	0.9292	0.8866	0.1

Examine Code



Next Step

Write Code to build, train,
evaluate, and use ML model