Machine Learning

Simplified for Developers with ML.NET

Virtual ML .NET Community Conference

Join the Conversation #virtualmlnet @jernej_kavka

V12

Jernej Kavka (JK)

SSW Senior Software Architect

@Jernej_kavka

github.com/jernejk

jkdev.me

linkedin.com/in/jernejkavka/

Brisbane AI User Group

.NET Core and EF Core dev by day

Cognitive Services and ML.NET enthusiast at night

Join the Conversation #virtualmlnet @jernej_kavka

.NET foundation

Cognitive Studio

Home page

Face API

Explore

Detect/Identify

Text API (beta)

Speech API (alpha)

Form Recognizer (alpha)

HOME › FACE › V1.0 › DETECT

Detect

Profiles
PoC

1. Get image

Image URL

File Upload

Web cam

2. Detect faces

Detection model
v02 (recommended)

More details

Recognition model (based on the g
v01 (group default)

More details

Detect (1 TR)

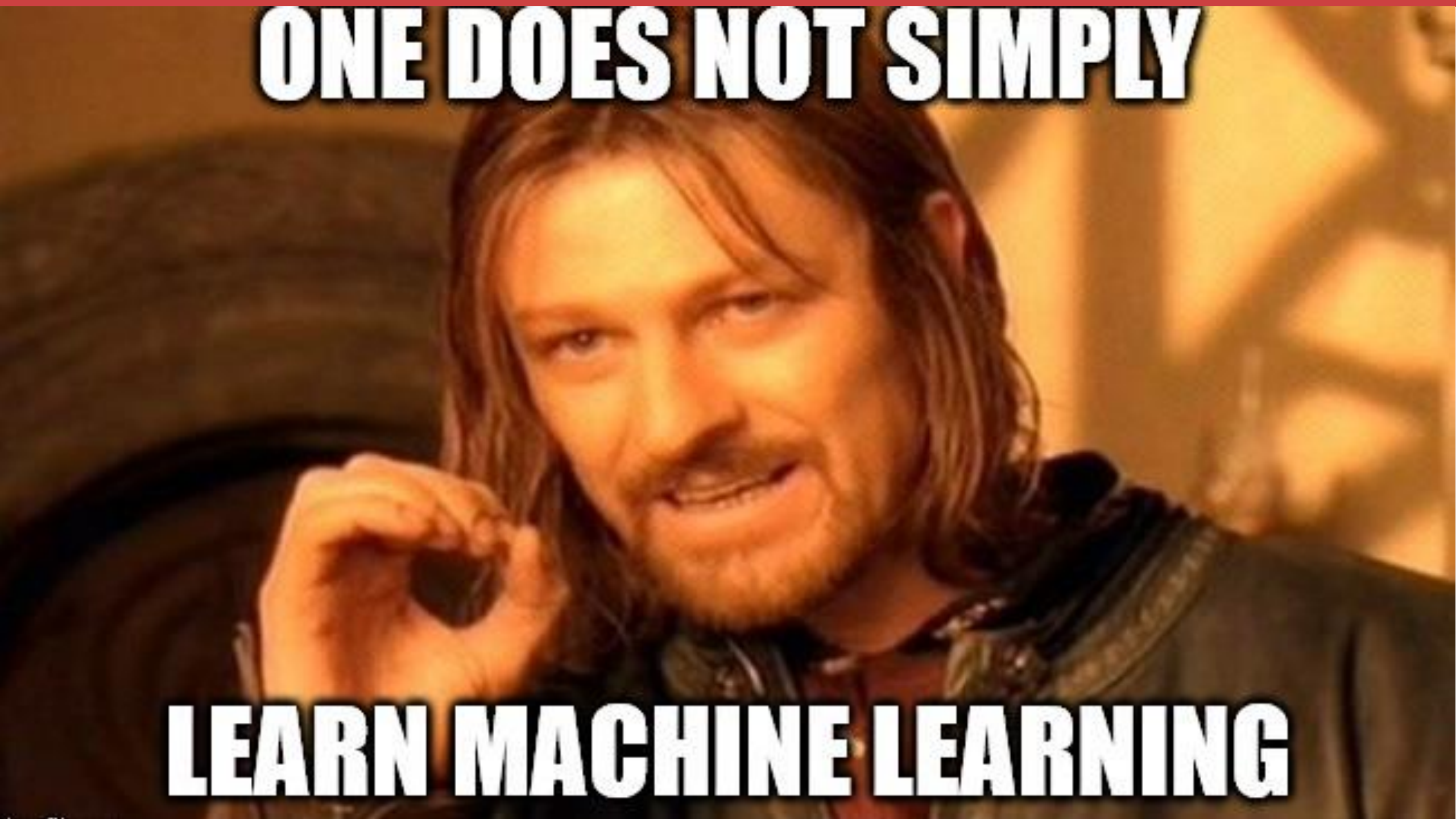
Cognitive Studio – https://cognitivestudio.dev

Table of Contents

My goal

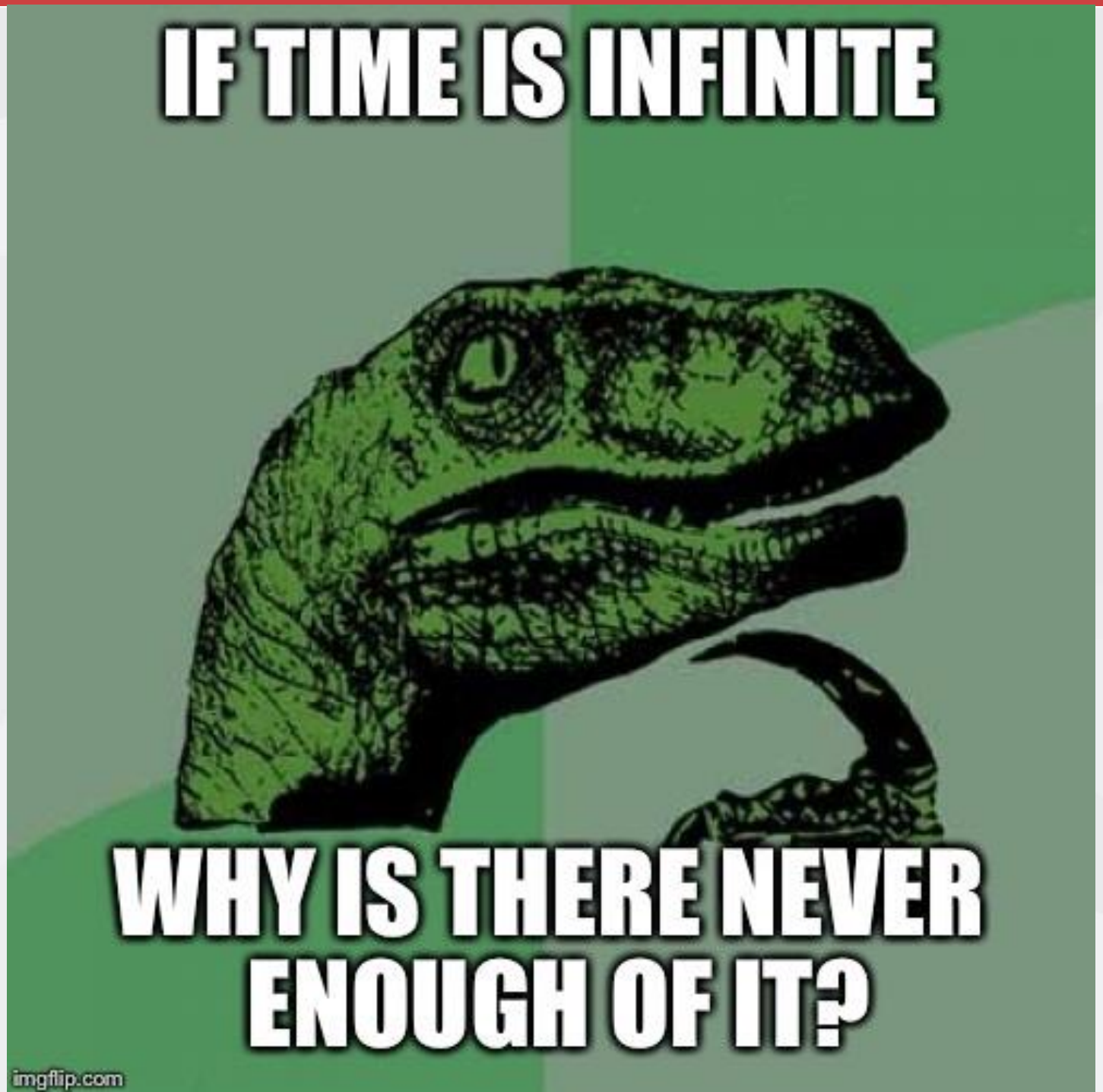
POS AUTHORISATION DOTNETFOUNDATION ORG REDMOND WA Card Used 0082	→	Investment
COFFEE 7 Melbourne	→	Food & Drink
Audible Australia Melbourne	→	Education


Microsoft
CNTK
POS AUTHORISATION
DOTNETFOUNDATION
ORG REDMOND WA Card Used 0082
COFFEE 7 Melbourne
Audible Australia Melbourne
Azure ML
Investment
Food & Drink
Education
TensorFlow


ONE DOES NOT SIMPLY
LEARN MACHINE LEARNING


TensorFlow


TensorFlow


IF TIME IS INFINITE
WHY IS THERE NEVER ENOUGH OF IT?
imgflip.com


FAILED DREAMS
FAILED DREAMS EVERYWHERE
memegenerator.net

ML.NET

An open source and cross-platform machine learning framework

Get started | Model Builder

Supported on Windows, Linux, and macOS

Built for .NET developers

With ML.NET, you can use your existing .NET skills to easily integrate ML into your .NET apps without any prior ML experience.

Custom ML made easy with AutoML

ML.NET offers AutoML and productive tools to help you easily build, train, and deploy high-quality custom ML models.

Extended with TensorFlow & more

ML.NET allows you to leverage other popular ML libraries like Infer.NET, TensorFlow, and ONNX for additional ML scenarios.

Trusted and proven at scale

Use the same ML framework used by recognized Microsoft products like PowerBI, Microsoft Defender, Outlook, and Bing.

Get started

Mode

Supported on Windows, Linux, and

Built for .NET developers

With ML.NET, you can use your existing .NET skills to easily integrate ML into your .NET apps without any prior ML experience.

Custom ML made easy with AutoML

ML.NET offers AutoML and productive tools to help you easily build, train, and deploy high-quality custom ML models.

ML.NET

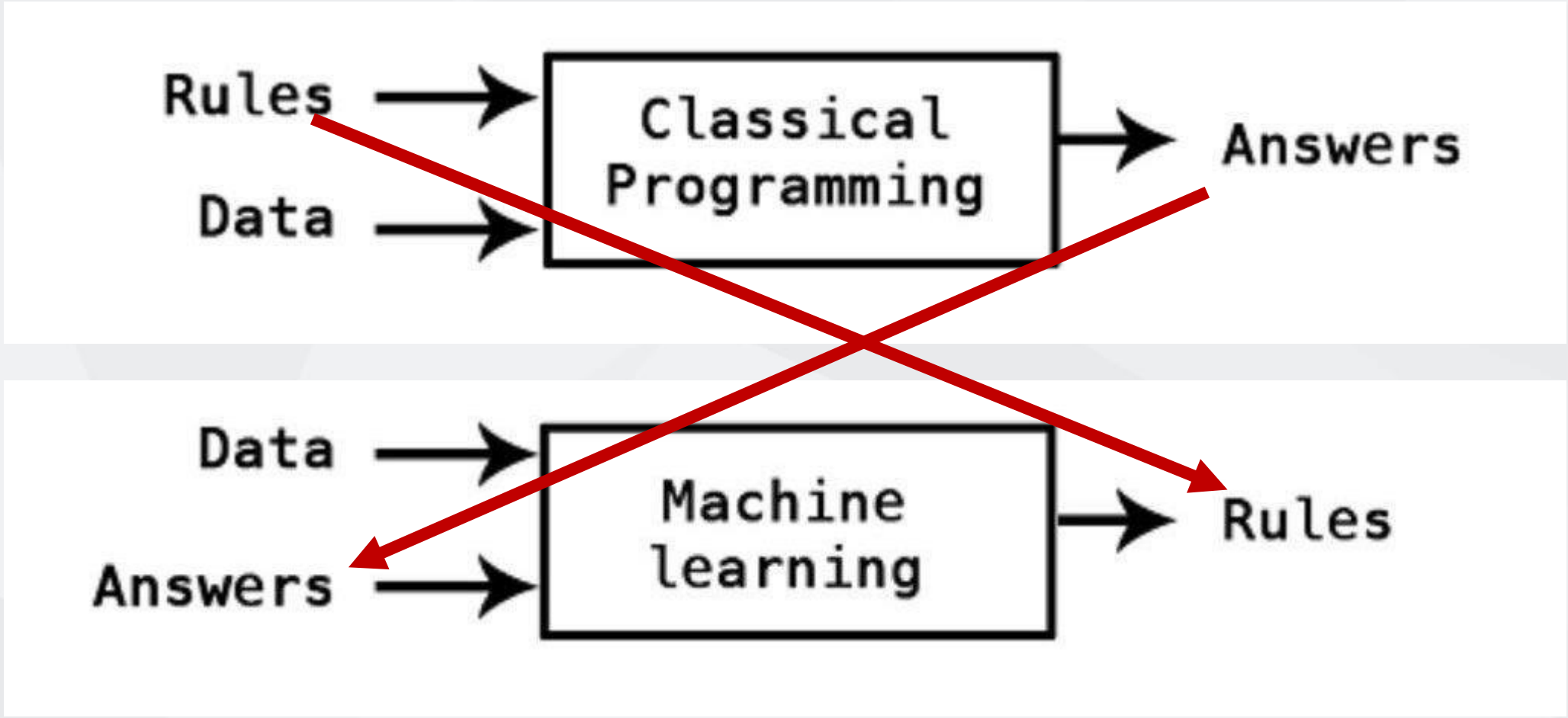
- MS machine learning SDK that works offline
- Simple yet powerful
- Already powering Microsoft products
- Support major ML models
  - TensorFlow
  - ONNX
- Awesome samples on GitHub
  - https://github.com/dotnet/machinelearning-samples

My goal

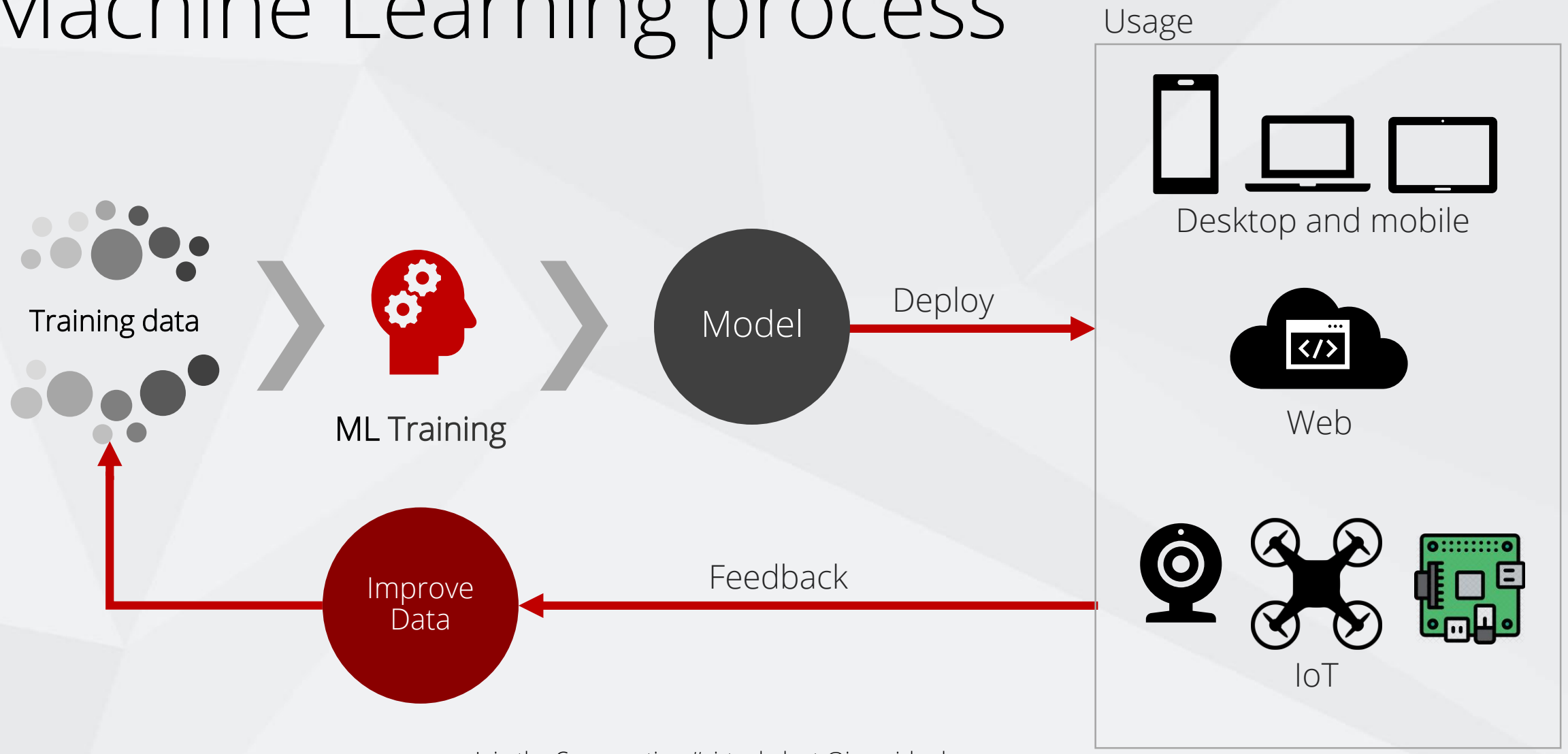
POS AUTHORISATION DOTNETFOUNDATION ORG REDMOND WA Card Used 0082	→	Investment
COFFEE 7 Melbourne	→	Food & Drink
Audible Australia Melbourne	→	Education

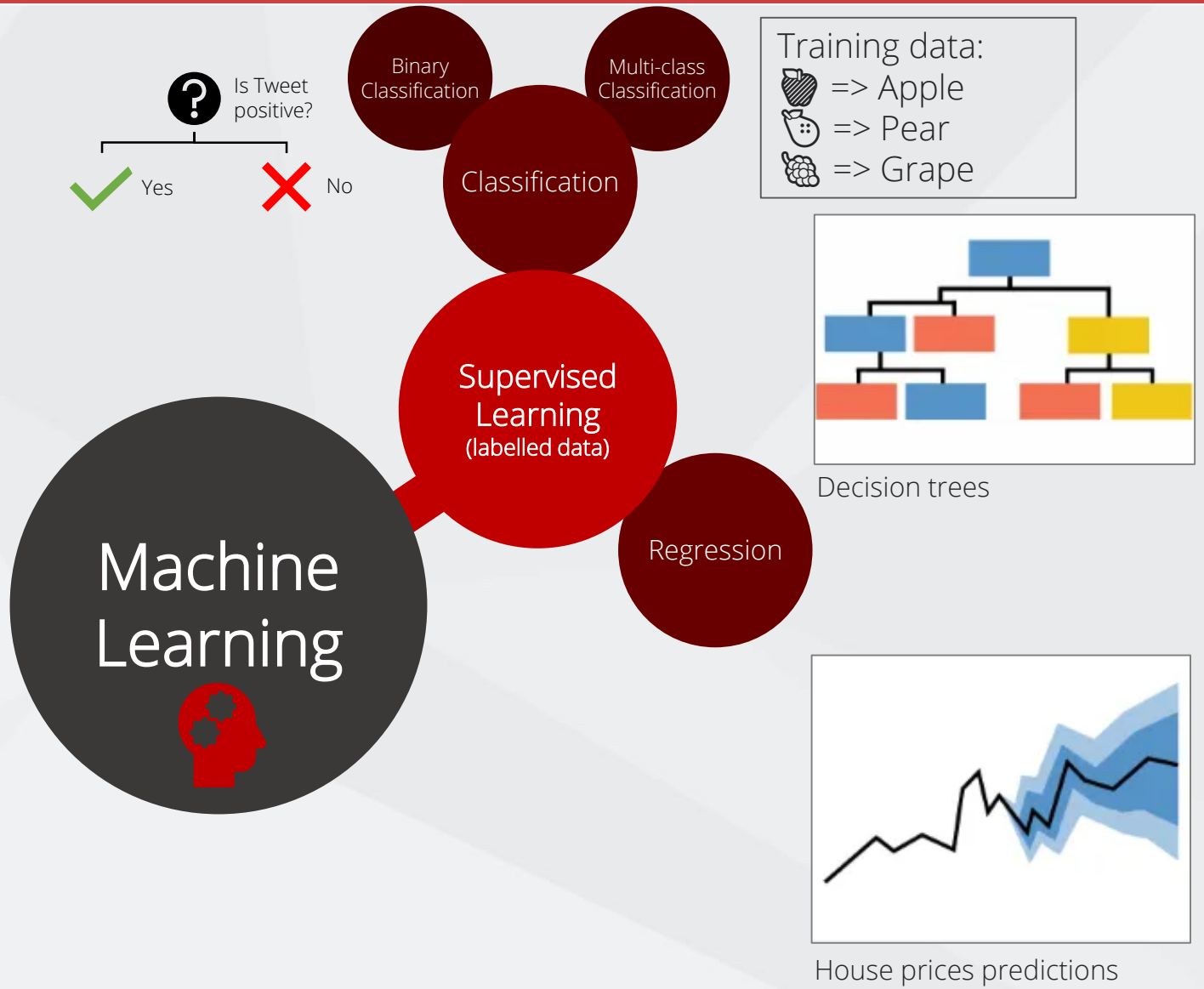
Table of Contents

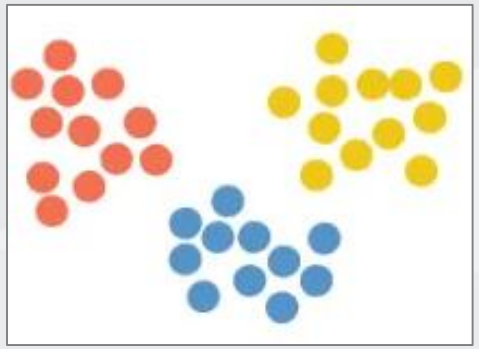
What is Machine Learning?



Machine Learning process




Is Tweet positive?
Yes
No
Binary Classification
Multi-class Classification
Classification
Supervised Learning
(labelled data)
Regression
Machine Learning
Training data:
=> Apple
=> Pear
=> Grape
Decision trees
House prices predictions



Product recommendations

Clustering

Unsupervised Learning (unlabelled data)

Dimension Reduction

Machine Learning

Supervised Learning (labelled data)

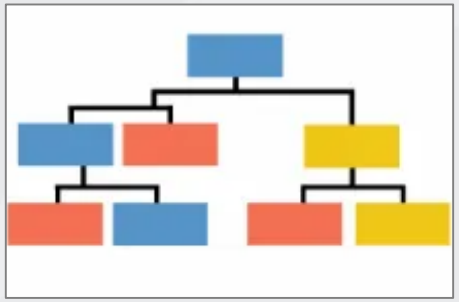
Classification

Binary Classification

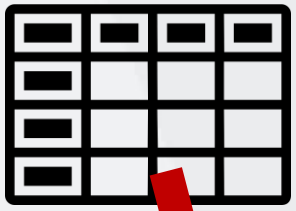
Multi-class Classification

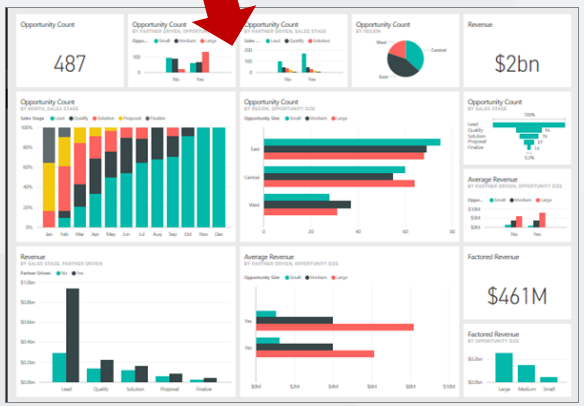
Regression

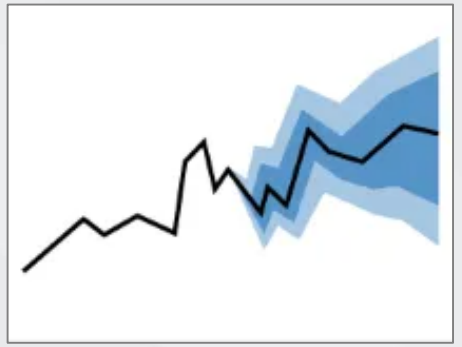
Training data:
=> Apple
=> Pear
=> Grape



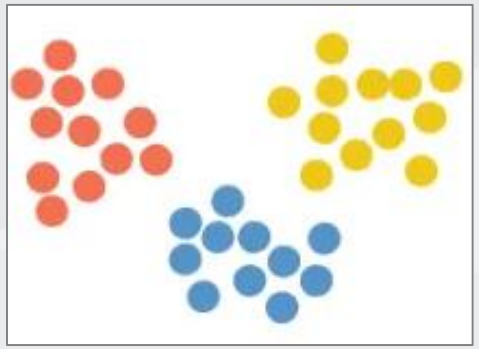
Decision trees







House prices predictions



Product recommendations

Clustering

Unsupervised Learning
(unlabelled data)

Dimension Reduction

Is Tweet positive?

Yes

No

Binary Classification

Multi-class Classification

Classification

Supervised Learning
(labelled data)

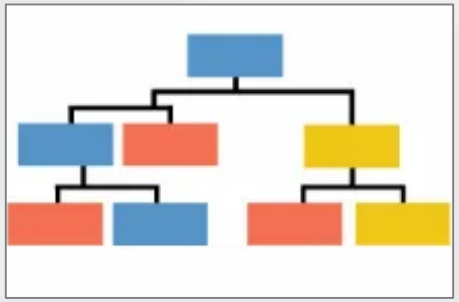
Regression

Training data:

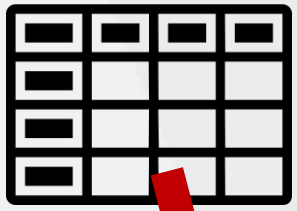
=> Apple

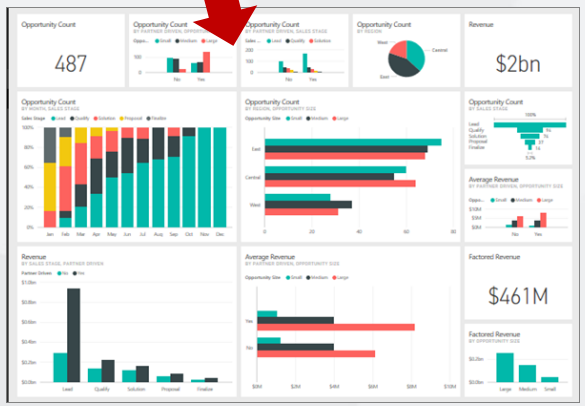
=> Pear

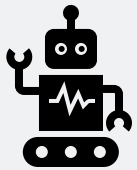
=> Grape



Machine Learning







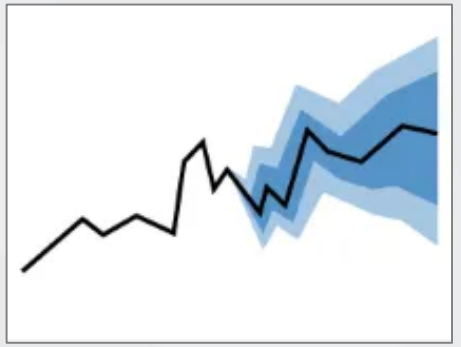
Reinforced Learning

Try

Fail

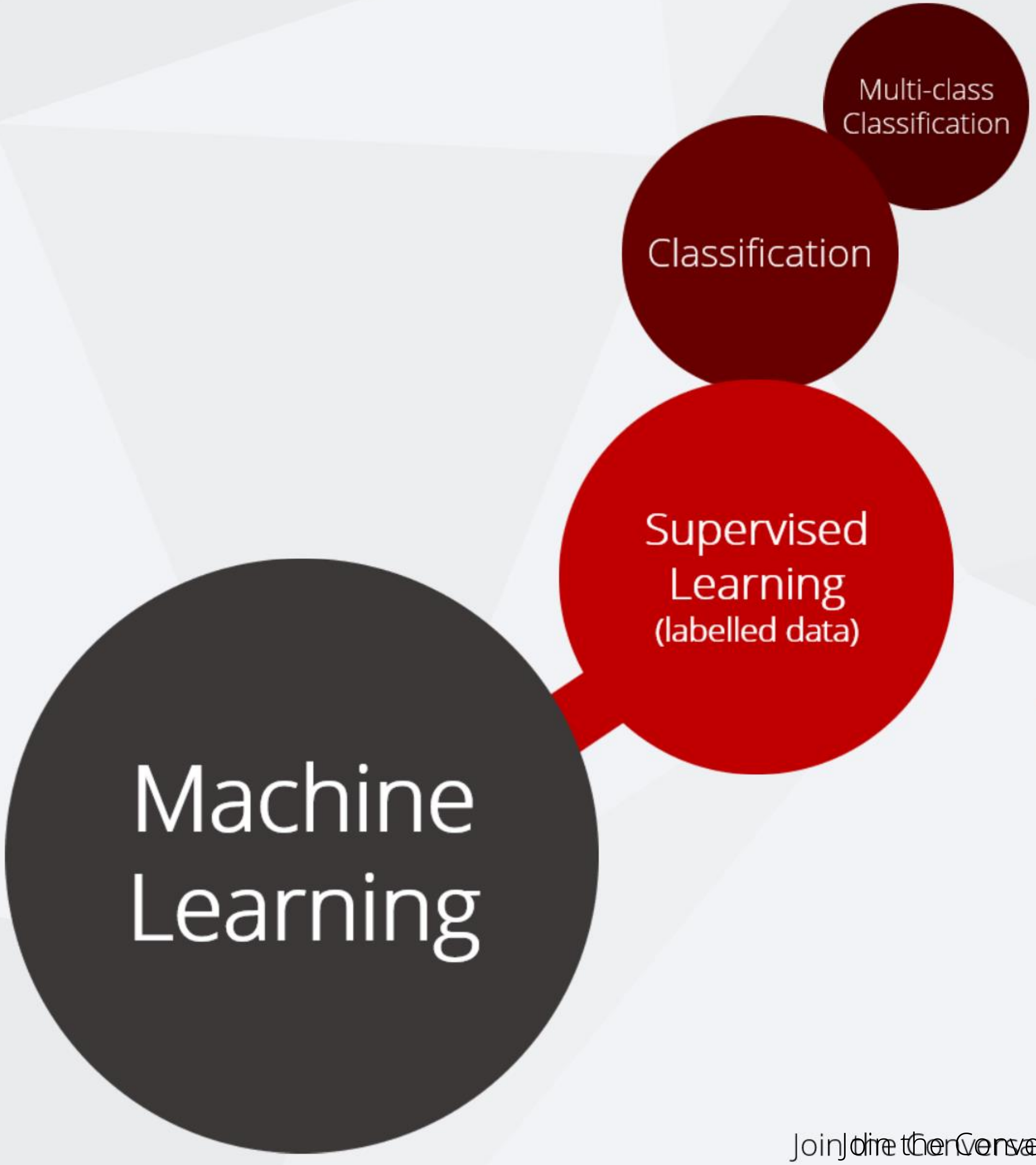
Learn

Repeat



House prices predictions

Join the Conversation #virtualmlnet @jernej_kavka



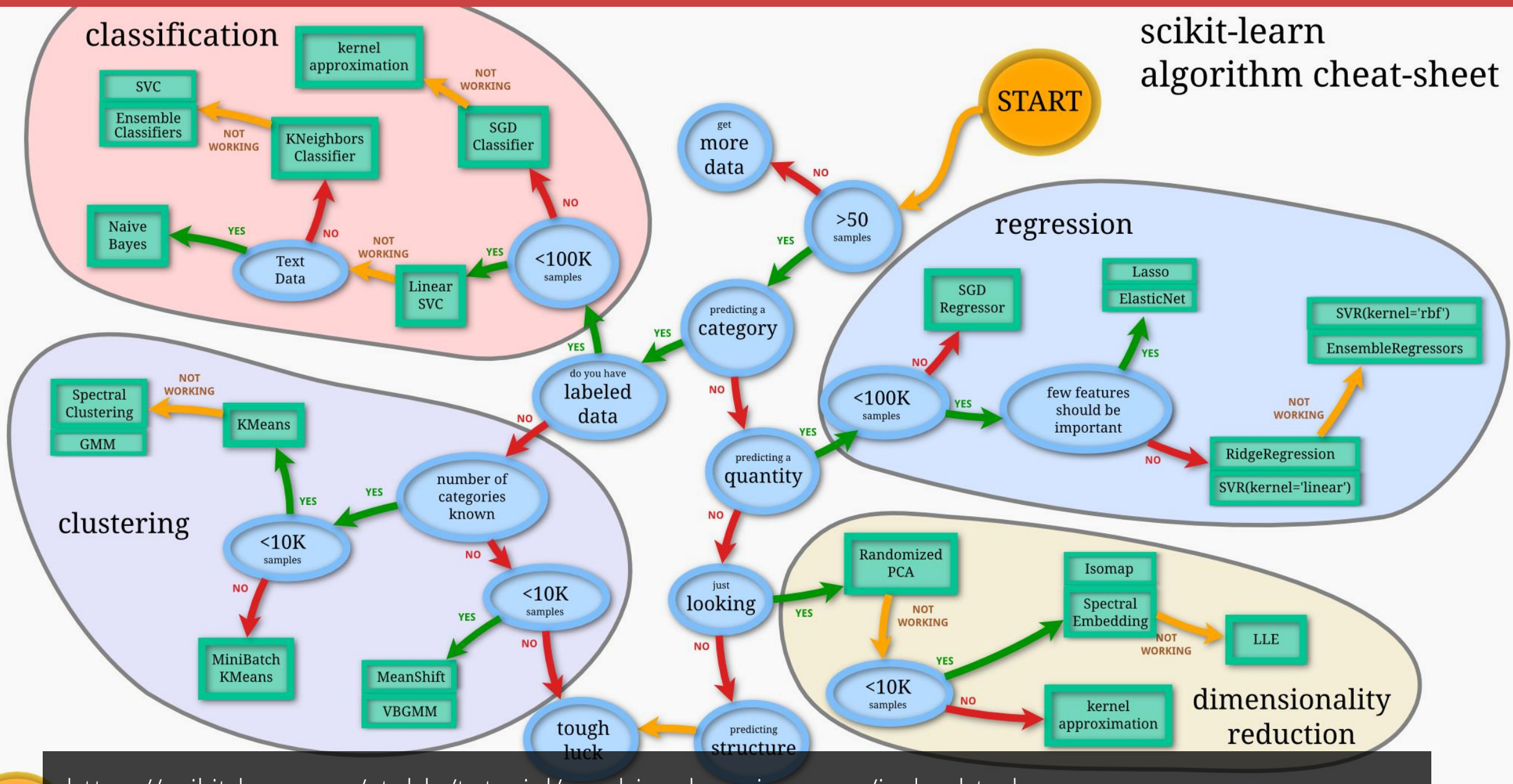
Multiclass classification trainers

You can train a multiclass classification model using the following training algorithms:

- LightGbmMulticlassTrainer
- SdcaMaximumEntropyMulticlassTrainer
- SdcaNonCalibratedMulticlassTrainer
- LbfgsMaximumEntropyMulticlassTrainer
- NaiveBayesMulticlassTrainer
- OneVersusAllTrainer
- PairwiseCouplingTrainer

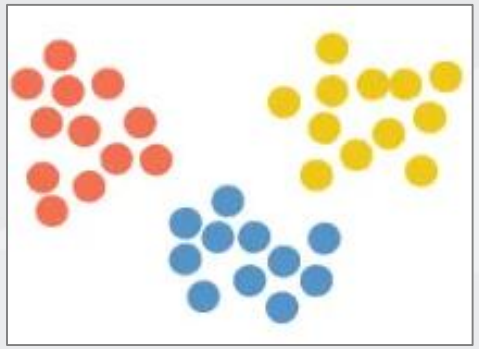
https://docs.microsoft.com/en-us/dotnet/machine-learning/resources/tasks

Join the Conversation #NDCPorto #MLNet @jernej_kavka



https://scikit-learn.org/stable/tutorial/machine_learning_map/index.html

Back



Product recommendations

Clustering

Unsupervised Learning
(unlabelled data)

Dimension Reduction

Is Tweet positive?

Yes

No

Binary Classification

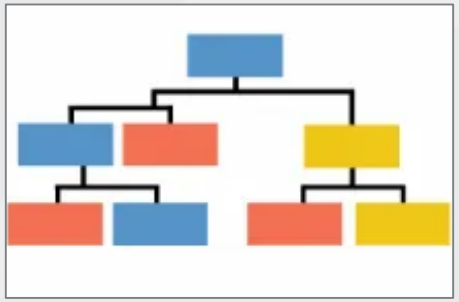
Multi-class Classification

Classification

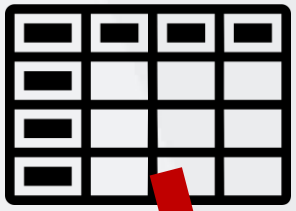
Supervised Learning
(labelled data)

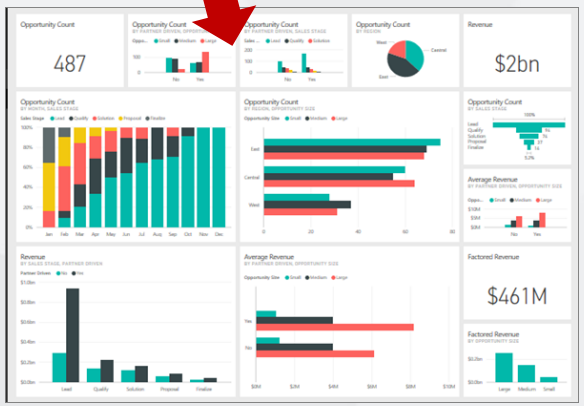
Regression

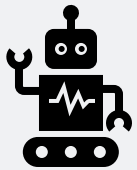
Training data:
=> Apple
=> Pear
=> Grape



ML.NET







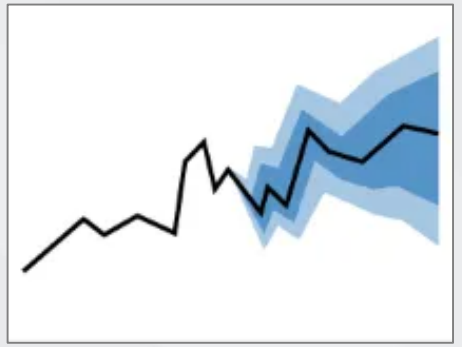
Reinforced Learning

Try

Fail

Learn

Repeat



House prices predictions

Join the Conversation #virtualmlnet @jernej_kavka

How do I start?

- Start with a scenario

Scenarios

- Is Tweet positive?

  Binary classification

- Categorize bank transactions

  Multi-class classification

- House price predictions

  Regression

- Product recommendation

  Clustering

Caveats

- ML won’t magically solve your problem
  - Understand your data
  - Guide ML toward right solution
  - Start with something simple like classification

Table of Contents

An open source and cross-platform machine lea

Get started | Model Builder

Supported on Windows, Linux, and macOS

Built for .NET developers

With ML.NET, you can use your existing .NET skills to easily integrate ML into your .NET apps without any prior ML experience.

Custom ML made easy with AutoML

ML.NET offers AutoML and productive tools to help you easily build, train, and deploy high-quality custom ML models.

Extended with TensorFlow &

ML.NET allows you to l
other popular ML librar
Infer.NET, TensorFlow, a
for additional ML scena

ML.NET scenarios

Sentiment analysis

Analyze the sentiment of customer reviews using a binary classification algorithm.

Product recommendation

Recommend products based on purchase history using a matrix factorization algorithm.

Price prediction

Predict taxi fares based on distance traveled etc. using a regression algorithm.

Customer segmentation

Identify groups of customers with similar profiles using a clustering algorithm.

Object detection

Recognize objects in an image using an ONNX deep learning model.

Fraud detection

Detect fraudulent credit card transactions using a binary classification algorithm.

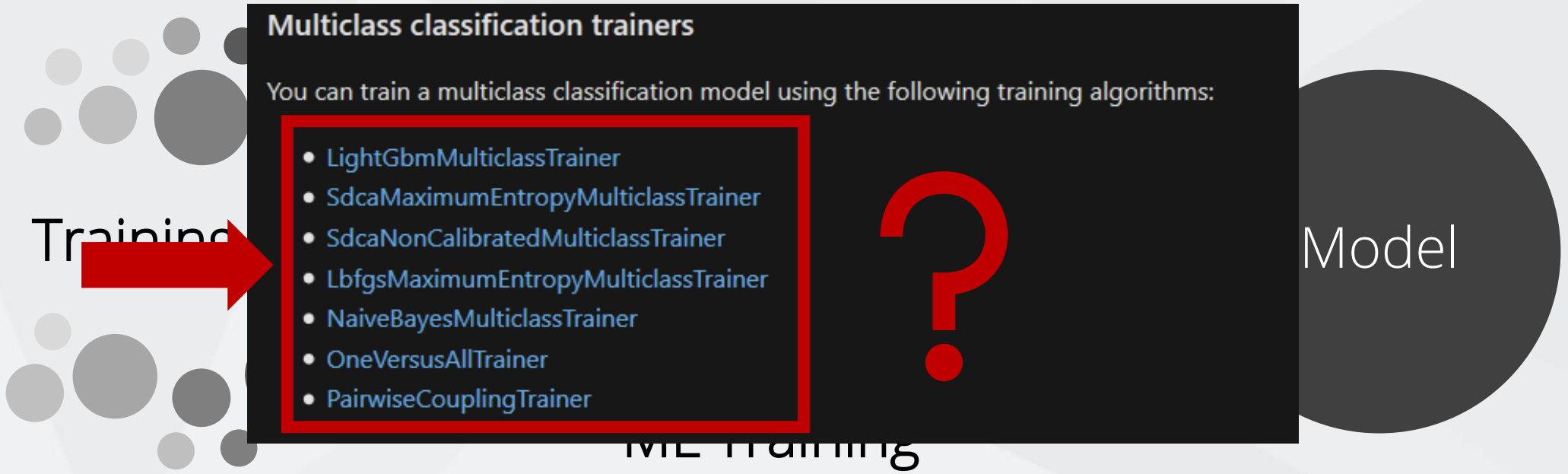
Sales spike

Image

Sales forecasting

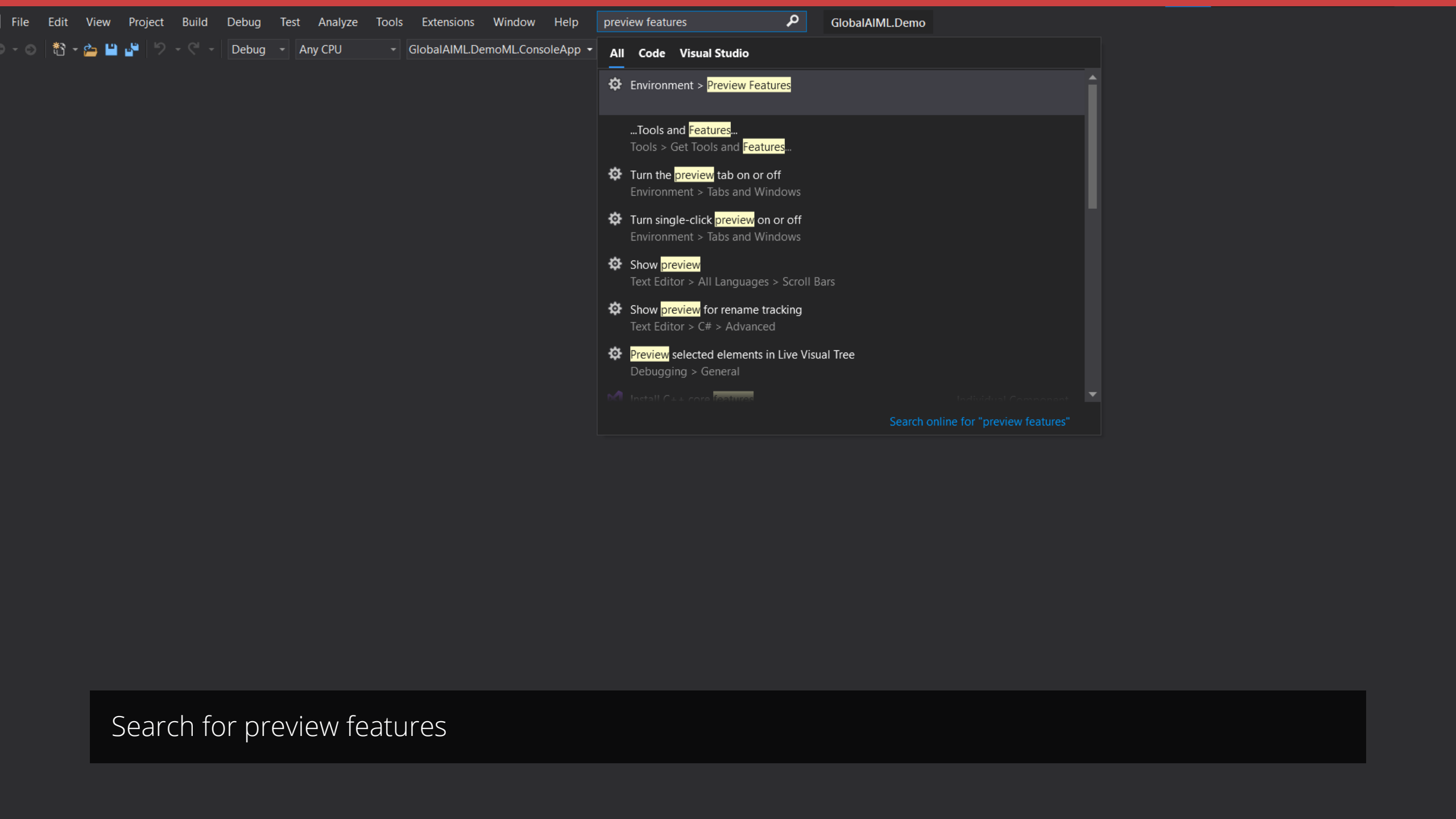
Build model (acquire knowledge)

Build model (acquire knowledge)

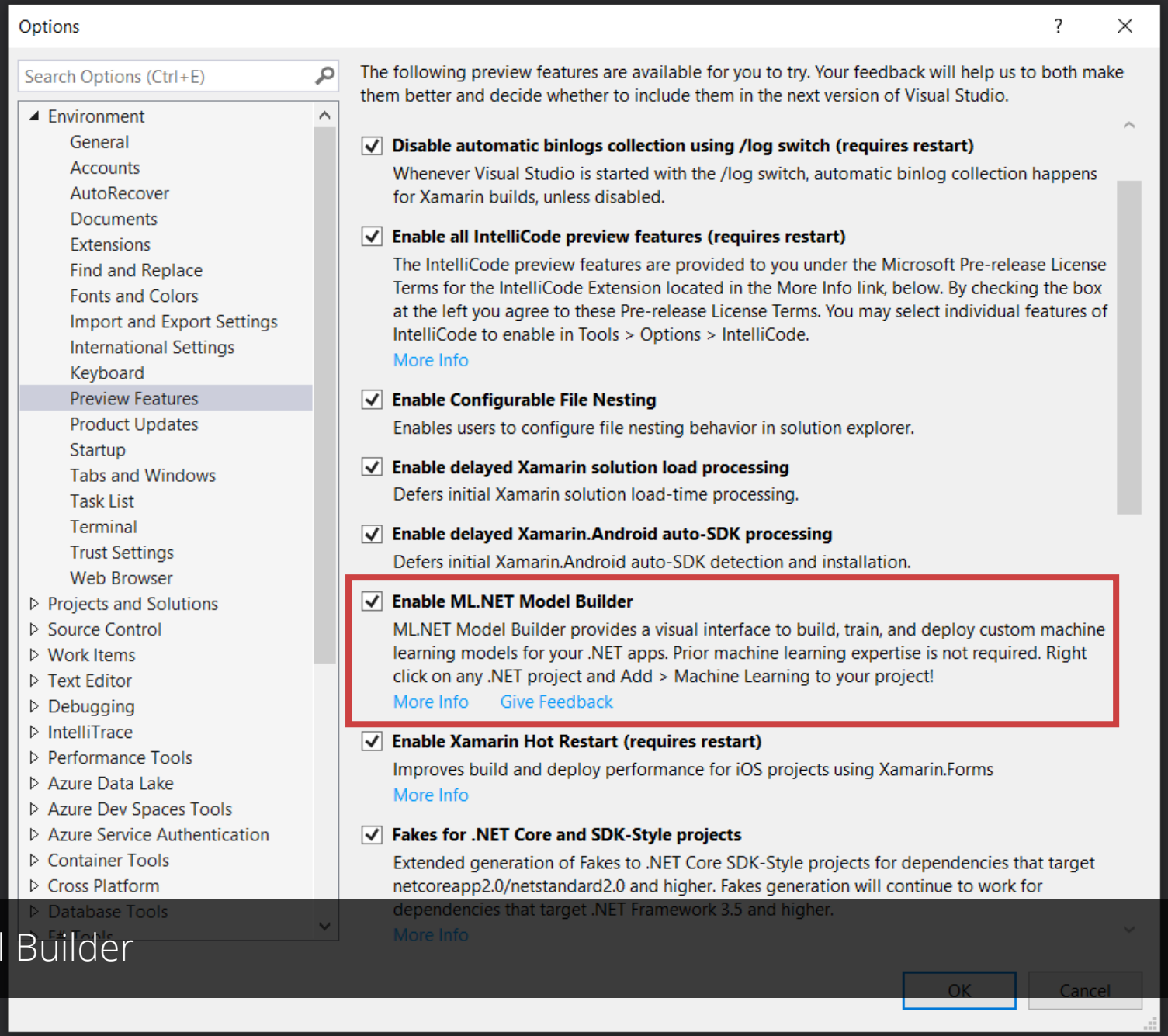


ML.NET Model Builder

- Wizard built on top of ML.NET SDK
- Picks best trainer for given scenario and data
  - Can be different one as the data changes over time
- Allows quick prototyping



Search for preview features

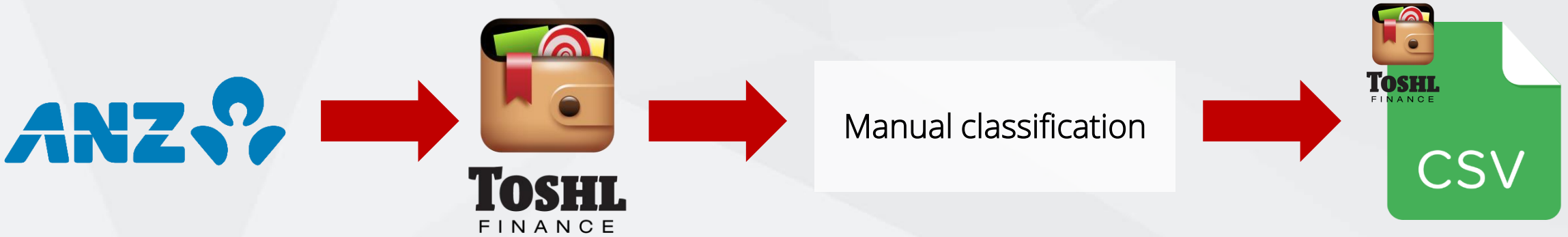


Enable ML.NET Model Builder

Let's do it

POS AUTHORISATION DOTNETFOUNDATION ORG REDMOND WA Card Used 0082	→	Investment
COFFEE 7 Melbourne	→	Food & Drink
Audible Australia Melbourne	→	Education

Training data preparation



Training data (CSV)

Date	Account	Category	Tags	Expense ar	Income ar	Currency	In main cur	Main curre	Description
17/05/2018	ANZ Rewards	Leisure	movies & music	11.99	0	AUD	11.99	AUD	Spotify P05F97AC4B Sydney
17/05/2018	Online Shopping	Food & Drinks	restaurants	15.15	0	AUD	15.15	AUD	VISA DEBIT PURCHASE CARD 0012 NAVALA CHURRASCARIA BRISBANE
17/05/2018	ANZ Rewards	Transport	taxi	23.25	0	AUD	23.25	AUD	PAYPAL *UBER AU 4029357733
17/05/2018	Online Shopping	Food & Drinks	coffee & tea	5.5	0	AUD	5.5	AUD	VISA DEBIT PURCHASE CARD 0012 AMERICAN CONCEPTS PT BRISBANE
17/05/2018	Online Shopping	Food & Drinks	bar	7	0	AUD	7	AUD	VISA DEBIT PURCHASE CARD 0012 RIVERLAND BRISBANE BRISBANE CITY
17/05/2018	Online Shopping	Food & Drinks	coffee & tea	24.91	0	AUD	24.91	AUD	VISA DEBIT PURCHASE CARD 0012 AT BROTHERS PTY LTD BRISBANE CITY
17/05/2018	Online Shopping	Food & Drinks	groceries	19	0	AUD	19	AUD	VISA DEBIT PURCHASE CARD 0012 RAW GROUP STAFFORD P STAFFORD
17/05/2018	Online Shopping	Leisure	movies & music	17.25	0	AUD	17.25	AUD	VISA DEBIT PURCHASE CARD 0012 THE HOYTS CORPORATIO STAFFORD
17/05/2018	Online Shopping	Food & Drinks	groceries	9.5	0	AUD	9.5	AUD	VISA DEBIT PURCHASE CARD 0012 RAW GROUP STAFFORD P STAFFORD
17/05/2018	Online Shopping	Food & Drinks	restaurants	15.15	0	AUD	15.15	AUD	VISA DEBIT PURCHASE CARD 0012 NAVALA CHURRASCARIA BRISBANE

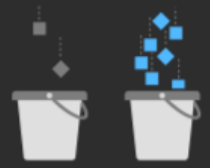
1. Scenario
2. Environment
3. Data
4. Train
5. Evaluate
6. Code

Select a scenario

Train with your data

The following scenarios use Automated ML to train and pick the best model for your data.

Learn more about training with your own data in Model Builder.



Text classification

Classify text data into 2+ categories, e.g. predict if comments are positive or negative sentiments.

Local ML

Value prediction

Predict a numeric value from your data (regression), e.g. predict the price of a house based on features like size, location, etc.

Local ML

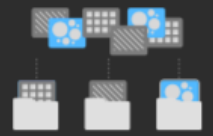


Image classification

Classify images into 2+ categories, e.g. predict whether an image is of a dog or a cat.

Azure ML Local ML

Recommendation

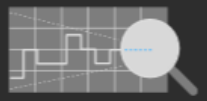
Produce a list of suggested items for a particular user, e.g. recommend products.

Local ML

Limited scenarios

The following scenarios are not yet supported by Automated ML, so walkthroughs with an example dataset and pre-defined training code are provided.

Learn more about examples in Model Builder.



Anomaly detection

Detect abnormalities or outliers in data. This example detects spikes in shampoo sales.

Forecasting

Predict future values based on previously observed time series values. This example predicts bike rental service demand.

Object detection

Detect and identify objects in images. This example detects objects (such as boats, people, and sofas) in images and draws

Clustering

Identify groups of related items without any pre-existing labels or categories. This example divides a set of iris flowers into different

1. Scenario
2. Environment
3. Data
4. Train
5. Evaluate
6. Code

Select training environment

This scenario only supports a local training environment.

Local
Train locally on your machine.

Local machine configuration

Processor	AMD Ryzen 5 3600 6-Core Processor
Memory	33GB

Next Step: get your data

Data

1. Scenario
2. Environment
3. Data
4. Train
5. Evaluate
6. Code

Add data

In order to build a model, you must add data and choose your column to predict.
How do I get sample datasets and learn more?

Input

Choose input data source from either SQL Server or File:

File

Select a file: ...

Supported file formats: .csv, .tsv or .txt.

Column to predict (Label): Select column

Input Columns (Features): Select column(s)

Data PreviewData preview

Select data to see the preview.

1. Scenario
2. Environment
3. Data
4. Train
5. Evaluate
6. Code

Add data

In order to build a model, you must add data and choose your column to predict.
How do I get sample datasets and learn more?

Input

Choose input data source from either SQL Server or File:

File

Select a file: C:\DataJK\ML\from-bank-unfiltere ...

Supported file formats: .csv, .tsv or .txt.

Column to predict (Label): Category

Input Columns (Features): 1 of 8 columns selected

Data PreviewData preview

10 of 2,553 rows and 2 of 9 columns. (1 Label, 1 Feature).

Category (Label)	Description
Leisure	Spotify P05F97AC4B Sydney
Other	AMAZON MKTPLC AU SYDNEY SOUTH
Transport	PAYPAL *UBER AU 4029357733
Transport	PAYPAL *UBER AU 4029357733
Other	AMAZON MKTPLC AU SYDNEY SOUTH
Food & Drinks	VISA DEBIT PURCHASE CARD 0012 AT BROTHERS PTY LTD BRISBANE CITY
Food & Drinks	VISA DEBIT PURCHASE CARD 0012 RAW GROUP STAFFORD P STAFFORD
Leisure	VISA DEBIT PURCHASE CARD 0012 THE HOYTS CORPORATIO STAFFORD
Food & Drinks	VISA DEBIT PURCHASE CARD 0012 RAW GROUP STAFFORD P STAFFORD
Food & Drinks	VISA DEBIT PURCHASE CARD 0012 NAVALA CHURRASCARIA BRISBANE

Next Step: train your model

Train

1. Scenario
2. Environment
3. Data
4. Train
5. Evaluate
6. Code

Train

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

Input

Time to train (seconds): 10

Start training

Progress

Start training to see progress and results.

1. Scenario
2. Environment
3. Data
4. Train
5. Evaluate
6. Code

Train

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

Input

Time to train (seconds): 10

Cancel training

Progress

Status: 2 seconds remaining...
Best accuracy: 84.71%
Best algorithm: AveragedPerceptronOva

1. Scenario
2. Environment
3. Data
4. Train
5. Evaluate
6. Code

Train

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

Input

Time to train (seconds): 60

Start training

Progress

Status: Training complete
Best accuracy: 84.53%
Best algorithm: AveragedPerceptronOva

Next step: evaluate your model

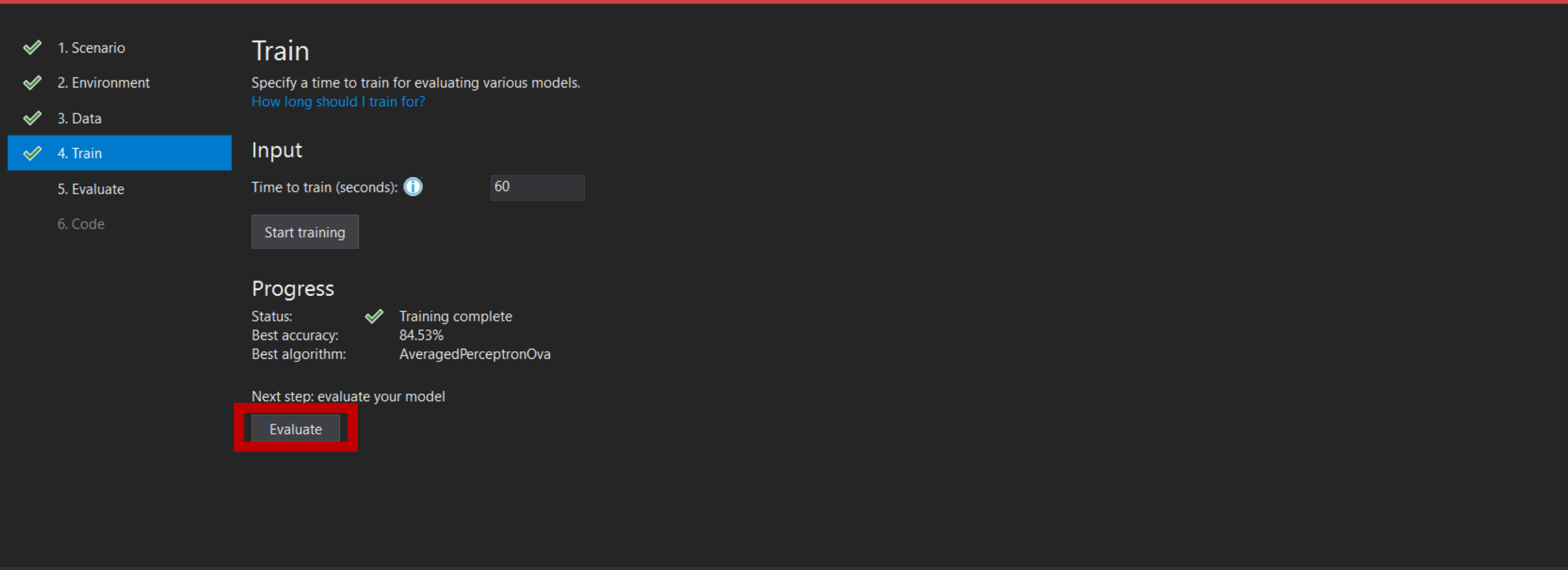
Evaluate

Output

Show output from: Machine Learning

```
|Total experiment time : 49.5641089 Secs                                                                                |
|Total number of models explored: 6                                                                                     |
-------------------------------------------------------------------------------------------------------------------------

|                                          Top 5 models explored                                                        |
-------------------------------------------------------------------------------------------------------------------------
|     Trainer                              MicroAccuracy  MacroAccuracy  Duration #Iteration                          |
|1    AveragedPerceptronOva                       0.8453         0.7189       6.1          1                          |
|2    SdcaMaximumEntropyMulti                     0.8430         0.6074       5.0          2                          |
|3    LinearSvmOva                                0.8264         0.6760       3.6          3                          |
|4    SymbolicSgdLogisticRegressionOva            0.8175         0.5084       6.3          4                          |
|5    LightGbmMulti                               0.6615         0.3155       5.3          5                          |
-------------------------------------------------------------------------------------------------------------------------
```



Output

Show output from: Machine Learning

```
|Total experiment time : 49.5641089 Secs                                                                                |
|Total number of models explored: 6                                                                                     |
-------------------------------------------------------------------------------------------------------------------------

|                                        Top 5 models explored                                                          |
-------------------------------------------------------------------------------------------------------------------------
|     Trainer                          MicroAccuracy  MacroAccuracy  Duration #Iteration                               |
|1    AveragedPerceptronOva                   0.8453         0.7189       6.1          1                               |
|2    SdcaMaximumEntropyMulti                 0.8430         0.6074       5.0          2                               |
|3    LinearSvmOva                            0.8264         0.6760       3.6          3                               |
|4    SymbolicSgdLogisticRegressionOva        0.8175         0.5084       6.3          4                               |
|5    LightGbmMulti                           0.6615         0.3155       5.3          5                               |
-------------------------------------------------------------------------------------------------------------------------
```

1. Scenario
2. Environment
3. Data
4. Train
5. Evaluate
6. Code

Evaluate

Results of training for your model can be found below. How do I understand my model performance?

Output

Overview | Details

ML task: multiclass-classification
Training time: 49.56 seconds
Models explored (total): 6 | View Top 5 models explored

Overall accuracy:

84.53%

Try your model

Sample data

Note: Fields below are pre-filled by row 1 of your data.

Description

Spotify P05F97AC4B Sydney

Top 5 Results:

Leisure	95%
Food & Drinks	2%
Home & Utilitie	1%
Other	1%
Transport	1%

Predict

Next step: Get your code snippet

Code

1. Scenario
2. Environment
3. Data
4. Train
5. Evaluate
6. Code

Code

Add the machine learning model and the projects and references for model consumption (**GlobalAIML.DemoML.Model**) and training/testing (**GlobalAIML.DemoML.ConsoleApp**) to your solution.

Add Projects

- ✓ 1. Scenario
- ✓ 2. Data
- ✓ 3. Train
- ✓ 4. Evaluate
- ✓ 5. Code

Code

Add the machine learning model and the projects and references for model consumption (**GlobalAIML.DemoML.Model**) and training/testing (**GlobalAIML.DemoML.ConsoleApp**)

Added Projects

Next Steps

1. Try the model

Run **GlobalAIML.DemoML.ConsoleApp** to make predictions on sample data.

2. Consume the model

In **GlobalAIML.Demo**, add the following using directive in the file where you will consume your model:

```
using GlobalAIML_DemoML.Model;
```

Then add your input data and use ConsumeModel.Predict() to load the model and predict the output:

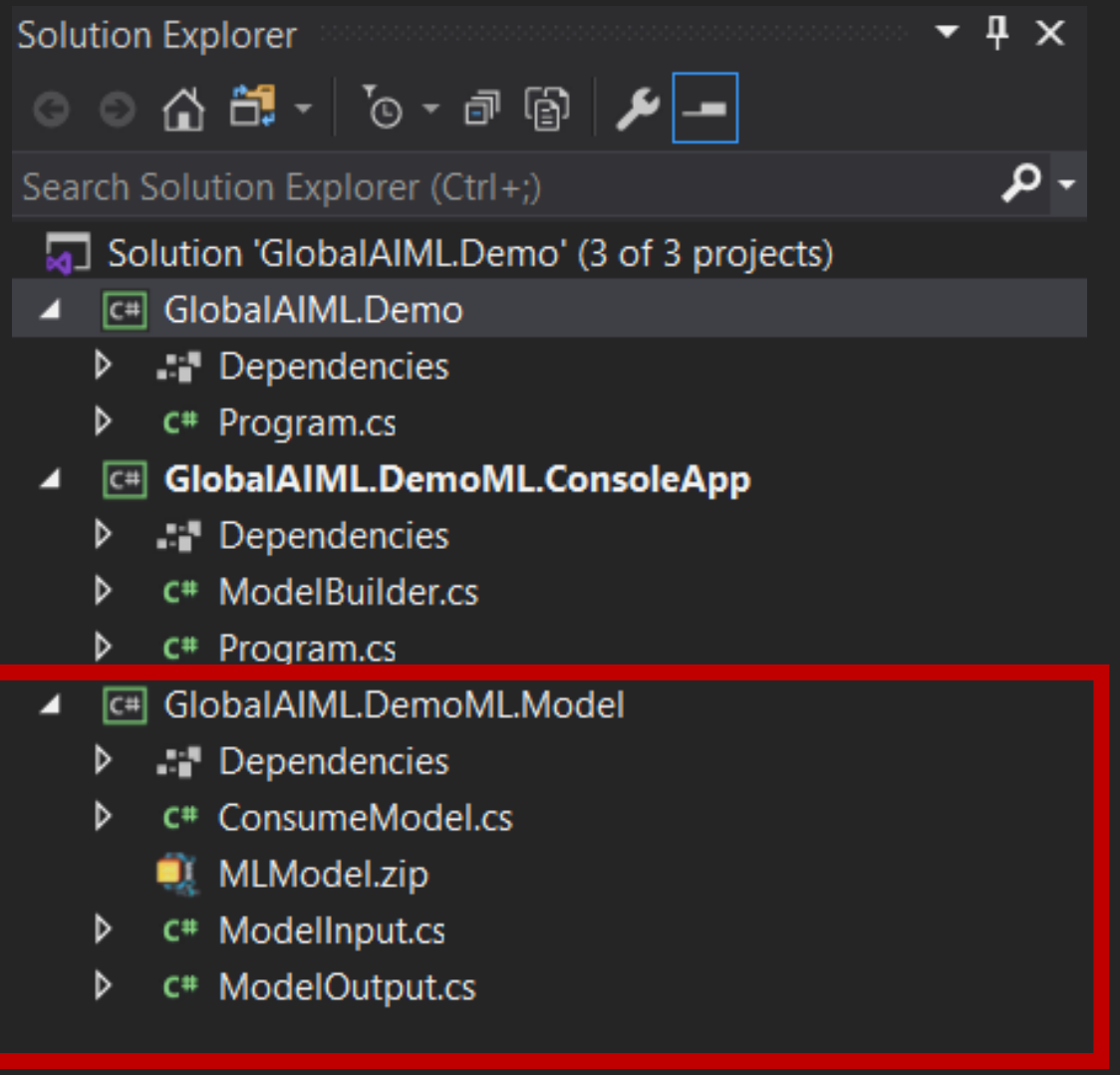
```
// Add input data
var input = new ModelInput();

// Load model and predict output of sample data
ModelOutput result = ConsumeModel.Predict(input);
```

Check out this tutorial to learn more about how to use the model in your app.

3. Improve the model

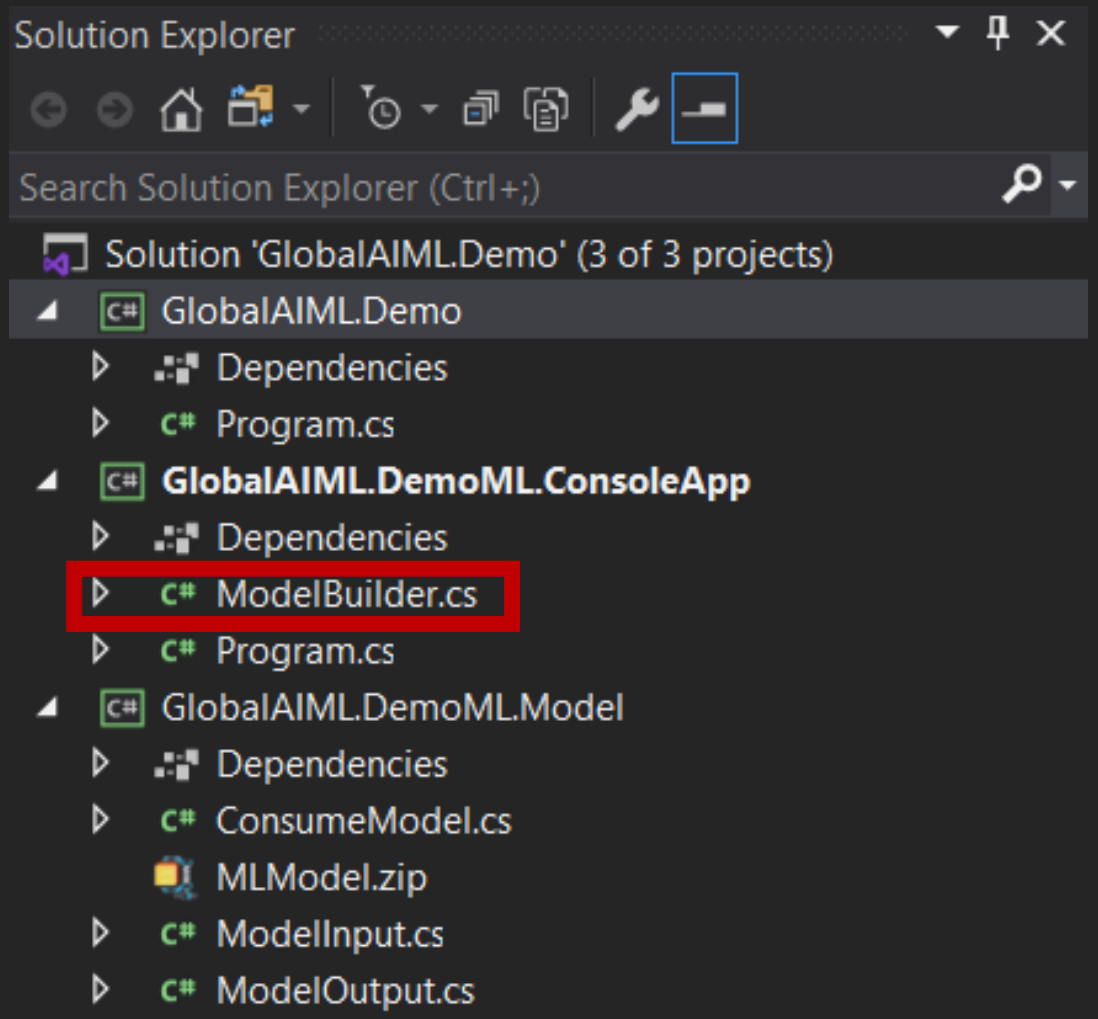
Train for a longer time, add more data, or learn more on the web.

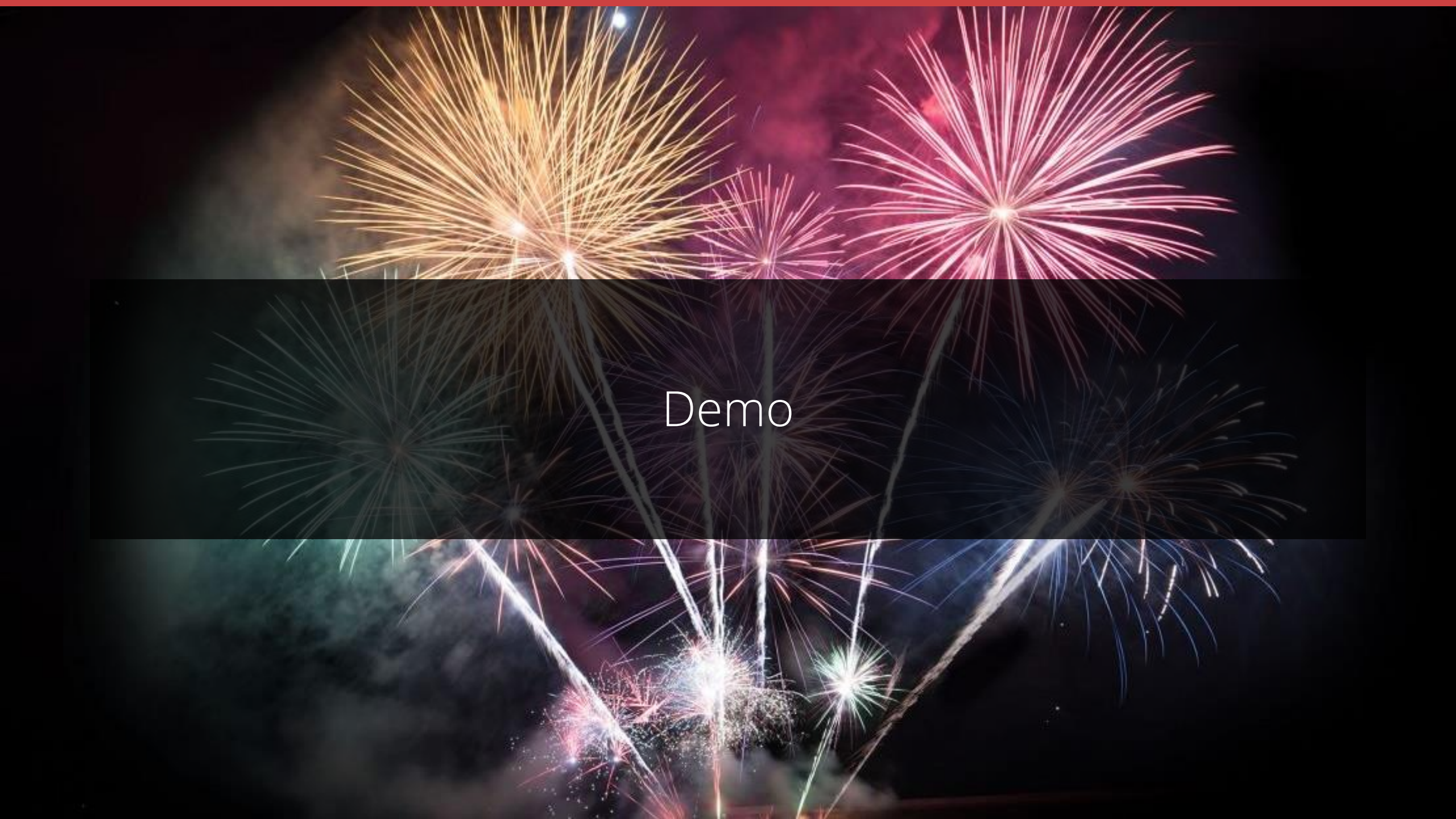

Solution Explorer
Search Solution Explorer (Ctrl+;)
Solution 'GlobalAIML.Demo' (3 of 3 projects)
GlobalAIML.Demo
Dependencies
Program.cs
GlobalAIML.DemoML.ConsoleApp
Dependencies
ModelBuilder.cs
Program.cs
GlobalAIML.DemoML.Model
Dependencies
ConsumeModel.cs
MLModel.zip
ModelInput.cs
ModelOutput.cs

Model


ML.NET

ML Training


Solution Explorer
Search Solution Explorer (Ctrl+;)
Solution 'GlobalAIML.Demo' (3 of 3 projects)
GlobalAIML.Demo
Dependencies
Program.cs
GlobalAIML.DemoML.ConsoleApp
Dependencies
ModelBuilder.cs
Program.cs
GlobalAIML.DemoML.Model
Dependencies
ConsumeModel.cs
MLModel.zip
ModelInput.cs
ModelOutput.cs


Demo

Summary

- Import CSV file
- Tweak columns to get better results
- Generate project with ML
- Tweaking generated code
- Minor issues with small datasets (100+ rows)

More details - http://bit.ly/jk-ml-interview-blog

Goal achieved

POS AUTHORISATION DOTNETFOUNDATION ORG REDMOND WA Card Used 0082	→	Investment
COFFEE 7 Melbourne	→	Food & Drink
Audible Australia Melbourne	→	Education

Thank you!

@Jernej_kavka

http://bit.ly/jk-ml-interview-blog

info@ssw.com.au

www.ssw.com.au

Sydney | Melbourne | Brisbane | Canberra